# Development of tools for genetic manipulation of *Staphylococcus aureus*

**Lúcia Branco Serra**

Dissertação para obtenção do

Grau de Mestre em Microbiologia Médica

**Dezembro, 2017**

# Development of tools for genetic manipulation of *Staphylococcus aureus*

**Lúcia Branco Serra**

Dissertação para obtenção do

Grau de Mestre em Microbiologia Médica

Orientador: Mariana Pinho, PhD, ITQB-UNL
Co-orientador: Helena Veiga, PhD, ITQB-UNL

**Instituto de Tecnologia Química e Biológica | Universidade Nova de Lisboa**

**Dezembro, 2017**

**©2017 Lúcia Serra**

O Instituto de Tecnologia Química e Biológica António Xavier e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

**Agradecimentos**

Este projecto representou um dos maiores desafios que até agora tive de enfrentar, e o seu sucesso não teria sido possível sem a contribuição de várias pessoas às quais gostaria de agradecer. Em retrospectiva, sinto-me muito afortunada de ter podido trabalhar neste Laboratório assim como de ter presenciado o seu bom ambiente de companheirismo, entreajuda e sobretudo de exigência científica, que muito me ajudou a ultrapassar esta etapa.

Gostaria primeiro de agradecer à minha orientadora Doutora Mariana Pinho, por dar-me a oportunidade de aprender no seu Laboratório e por transmitir o seu conhecimento sempre que precisei. O seu acompanhamento e sentido crítico foram fundamentais para mim e ensinaram-me muito mais do que até agora tinha aprendido.
À minha co-orientadora, a Doutora Helena Veiga, que esteve sempre disponível para mim e me motivou a apresentar a "melhor tese do mundo". O desenvolvimento deste projecto não teria sido possível sem o seu apoio, que me mostrou a responsabilidade que temos como cientistas de apresentar trabalhos bem-feitos e exigentes.

Um agradecimento sincero a ambas por me terem transmitido o seu gosto genuíno pela Ciência.

Gostaria igualmente de agradecer ao Doutor Marc Bramkamp, por ter cedido os plasmídeos pKILL, assim como todo o material e informação necessários para testá-los neste projecto, provando ser ferramentas com muito potencial.

Não podia escrever este texto sem deixar um "obrigada" sentido a todas as pessoas do Laboratório (MGPI/MGPII e SF), que além de aguentarem as minhas perguntas incessantes e os meus almoços muuuito lentos, sempre me deram o seu apoio e atenção. O bom ambiente que proporcionaram (incluindo deliciosas partilhas de chocolates), assim como as sugestões e os ensinamentos dentro e fora da ciência, foram extremamente valiosos para mim e sem eles, a minha estadia neste Laboratório não teria tido a mesma importância. Cada um de vocês me ajudou no neste projecto e portanto, as minhas realizações são também vossas.

Para terminar, queria agradecer aos meus amigos, que mesmo perto ou longe, biólogos ou não-biólogos, portugueses ou não-portugueses, me ajudaram neste percurso nas mais diversas formas. A vossa presença contribuiu enormemente tanto para a

manutenção da minha sanidade mental nesta fase como para o meu crescimento pessoal. Sem vocês este caminho teria sido sem dúvida muito mais difícil e aborrecido…

Gostaria finalmente de agradecer aos meus pais, a quem dedico esta tese, pelo seu apoio incondicional e por incentivar-me desde sempre a ser melhor e a ambicionar o melhor. A vossa compreensão e sobretudo a vossa presença incessante foram muito importantes para mim nesta etapa e ajudaram-me muito mais do que alguma vez poderei agradecer. Vocês são o meu exemplo. Obrigada.

Este texto não foi escrito sob o Acordo Ortográfico.

# Abstract

With the global distribution of multidrug resistant *Staphyloccocus aureus,* a human pathogen responsible for thousands of deaths every year, there is an increasing need to study its mechanisms of resistance as well as discover new antibiotics that effectively interfere with its growth and division. However, despite the growing interest in studying this bacteria, there are only a few tools currently available for the genetic manipulation of *S. aureus*. In this study we developed genetic tools that could facilitate DNA insertion and gene repression in *S. aureus*: pKILL plasmid and CRISPRi.

The integrative pKILL plasmid was designed and constructed by M. Bramkamp to insert/delete a DNA fragment from *S. aureus* genome in a single step, by taking advantage of a double crossover event. Since it has already been successfully applied in *Bacillus subtilis*, and when functional, is less laborious and time-consuming than other currently used techniques, our objective was to test whether this plasmid could be used in *S. aureus*.

CRISPRi is a system designed to repress a DNA target by introducing a dead-Cas9 nuclease that specifically blocks its expression. CRISPRi has been more used in the genetic manipulation of eukaryotes than prokaryotes, even though CRISPR is a bacterial immunity system. Here, our goal was to develop a CRISPRi repression system, using a *S. aureus* dCas9 to achieve gene expression control in *S. aureus*, instead of the widely used dCas9 from *Streptococcus pyogenes*.

We have successfully used pKILL to genetically manipulate *S. aureus*, albeit at a very low efficiency, and implemented a CRISPRi system to repress non-essential *S. aureus* genes.

# Resumo

Com a distribuição global de *Staphylococcus aureus* multirresistente, um patogénio responsável por milhares de mortes cada ano, existe uma necessidade premente em estudar os seus mecanismos de resistência assim como de descobrir novos antibióticos que interfiram eficientemente com o seu crescimento e divisão. Contudo, apesar do crescente interesse em estudar esta bactéria, existem poucas alternativas atualmente disponíveis para a manipulação genética de *S. aureus*. Neste estudo desenvolvemos duas técnicas que poderiam facilitar a inserção de DNA e repressão génica em *S. aureus*: o plasmídeo pKILL e CRISPRi.

O plasmídeo pKILL integrativo foi construído por M. Bramkamp para inserir/remover um fragmento de DNA do genoma de *S. aureus* num único passo, através de um evento de recombinação dupla. Uma vez que já foi aplicado com sucesso em *Bacillus subtilis*, e, quando funcional, é menos trabalhoso e demorado do que outras técnicas atualmente utilizadas, o nosso objetivo foi testar se este plasmídeo poderia ser utilizado em *S. aureus*.

CRISPRi é um sistema desenhado para reprimir um alvo de DNA introduzindo uma nuclease dead-Cas9 que bloqueia especificamente a sua expressão. CRISPRi tem sido mais utilizado na manipulação genética de eucariontes que procariontes, apesar de CRISPR ser um sistema de imunidade bacteriano. Neste estudo, pretendemos desenvolver um sistema de repressão CRISPRi utilizando a dCas9 de *S. aureus* para conseguir o controlo da expressão génica de *S. aureus*, em vez da dCas9 de *Streptococcus pyogenes* mais frequentemente utilizada.

Implementámos com sucesso a utilização do plasmídeo pKILL, embora com muito baixa eficiência, e um sistema CRISPRi para reprimir genes não-essenciais de *S. aureus*.

# Keywords

*Staphylococcus aureus*; fluorescence microscopy; genetic manipulation; CRISPRi

# Contents

# List of figures

# List of tables

# List of abbreviations

**AAV** – Adeno-associated virus

**Amp -** Ampicillin

**asRNA** – antisense RNA

**CA-MRSA** – Community acquired methicillin resistant *Staphylococcus aureus*

**Cas** – CRISPR associated

**Cm -** Cloramphenicol

**CRISPR -** Clustered Regularly Interspaced Short Palindromic Repeats

**CRISPRi** – CRISPR interference

**crRNAs** – CRISPR RNA

**crRNP -** CRISPR ribonucleoprotein

**dCas9** – dead Cas9

**DSB** – Double strand break

**EDTA -** Ethylenediaminetetraacetic acid

**Ery -** Erythromycin

**GFP** – Green fluorescent protein

**gRNA** – guide RNA

**HA-MRSA** – Hospital acquired methicillin resistant *Staphylococcus aureus*

**IPTG -** Isopropyl β-D-1-thiogalactopyranoside

**Kan -** Kanamycin

**LA** – Luria Agar

**LB** – Luria Broth

**MOPS -** 3-(N-morpholino)propanesulfonic acid

**MRSA** – Methicillin resistant *Staphylococcus aureus*

**MSA** – Mannitol salt agar

**Neo** - Neomycin

**OD** – Optical density

**PAM** – Protospacer adjacent motif

**PCR** – Polymerase chain reaction

**PI** – PAM interacting

**pre-crRNA** – precursor CRISPR RNA

**RM** – Restriction-modification

**SaCas9** – *Staphylococcus aureus* Cas9

**sfGFP** – super fast GFP

**sgRNA** – single guide RNA

**SpCas9** – *Streptococcus pyogenes* Cas9

**TSA** – Tryptic soy agar

**TSB** – Tryptic soy broth

**tracrRNA** – trans-activating CRISPR RNA

**X-gal** - 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

# 1. Introduction

## A brief introduction to *Staphylococcus aureus*

*Staphylococcus aureus* is a Gram-positive bacteria with low G+C content, characterized by non-motile, non-spore forming individual cocci.[1] It was first described in 1880 by Alexander Ogston, who isolated these bacteria from surgical wounds. Their characteristic shape and color gave rise to the name by which the microorganisms are known today.[2,3]

Although it is a common commensal of the skin and nares (colonizing approximately 1/3 of the human population worldwide), it is also an important human pathogen, responsible for skin/soft tissue and device related infections, bacteremia and infective endocarditis.[2,4] It is normally transmitted by skin-to-skin contact between individuals or contaminated objects, infecting predominantly low birth-weight neonates and immunosuppressed patients requiring medical devices/procedures, and therefore being a major cause for nosocomial infections.[2,5–7] An annual mortality rate of more than 20000 is attributed to *S. aureus* infections in the United States, a higher value than AIDS (acquired immune deficiency syndrome) and tuberculosis combined in the same country.[8]

*S. aureus* infections are commonly treated with antibiotic chemotherapy. With the discovery of penicillin in 1929 by *Sir* Alexander Fleming, antibiotic chemotherapy launched a new era in the history of medicine, permitting for the first time an effective attack against bacterial infections and radically reducing the fatality rate for *S. aureus* related diseases.[4] However, along with the usage of antibiotics also came drug resistance, and already in 1948 it was estimated that 60% of *S. aureus* hospital isolates were resistant to penicillin.[9] Although new antibiotics have been developed afterwards to fight these infections, treatment efficacy has reached a plateau over the past several decades, as reflected by the ever-present numbers of *S. aureus* cases.[4]

Drug resistance was a consequence of the increasing use of antibiotics. The first *S. aureus* strains resistant to penicillin expressed a penicillinase/β-lactamase enzyme that degraded the β-lactam ring in the penicillin structure and inactivated its effect.[10] With the later application of different penicillinase-resistant β-lactams such as methicillin and oxacillin, *S. aureus* further increased its drug resistance by acquiring a specific penicillin-

binding protein, PBP2a, which is not inhibited by these antibiotics.[2] The presence of PBP2a enabled not only resistance to methicillin but also to other β-lactams, such as cephalosporines and carbapenems, thus representing the molecular basis for broad-spectrum resistance.[2,11]

The growing presence of methicillin resistant *S. aureus* (MRSA) since its first isolation in the UK, led to its spread all over the world in the years that followed, hence rending β-lactams virtually ineffective against *S. aureus* infections.[12,13] The introduction of new antibiotics later on led to the adaptation and selection of increasingly resistant *S. aureus* strains, which are now more commonly known as multidrug resistant *S. aureus* (MRSA). Nowadays, MRSA isolates account for more than 50% *S. aureus* infections in Portugal and more than 30% in the United States, and are normally found in health care settings (hospital acquired, HA-MRSA) as nosocomial pathogens.[2,14] However, with the recent emergence of MRSA infections in healthy individuals, acquired in the community (CA-MRSA), MRSA outbreaks have been spreading both inside and outside medical environments - limiting its control.[15] With such globally distributed HA-MRSA and CA-MRSA clones, the urgency for a better understanding of *S. aureus* resistance mechanisms and for developing new antibiotics is pressing.

## Genetic engineering in *S. aureus*

Genetic manipulation in *S. aureus* is currently achieved by different techniques. In this chapter we will summarize the ones that are more commonly used.

Shuttle *Escherichia coli/S. aureus* plasmids have a fundamental role in molecular cloning in *S. aureus* since they allow for a certain DNA of interest to be constructed in an *E. coli* strain and later to be introduced into *S. aureus*. They benefit from the accessible *E. coli* cells transformation and plasmid propagation before incorporation into Gram-positive bacteria. However, *S. aureus* has native restriction-modification (RM) systems that are major barriers against the uptake of foreign DNA, the basis of cellular genetic engineering. A number of laboratory strains have been developed in order to promote the uptake of shuttle plasmids without RM interference. Two of these strains are the cytosine methylation-deficient *E. coli* DC10B strain (originated from high efficiency cloning *E. coli* strain DH10B), and RN4220, a highly mutagenized *S. aureus* strain (originated from

strain 8325-4, cured of three prophages) that is able to accept foreign DNA and to methylate it as native, "disguising" its real origin.[16–18] RN4220 is extensively used as the initial *S. aureus* recipient strain, before transferring the foreign DNA to another *S. aureus* strain.[18] Electroporation and transduction are some of the methods of choice to introduce DNA fragments into *S. aureus* cells: electroporation enables the uptake of *E. coli* shuttle plasmid by creating pores into competent *S. aureus* cells by electric discharges; and transduction allows the transfer of desired DNA products between *S. aureus* cells by phage delivery.[16]

For gene expression control, there are a myriad of possibilities from which only a few are commonly used in *S. aureus*. We will summarize two of these approaches that are currently used: inducible promoters and RNA antisense technology.

Inducible promoters are able to directly alter the expression of proximal genes in the presence of an inducer, the molecule required to activate the promoter. Generally, inducible promoters are useful not only to prove gene essentiality (by preventing the expression of a certain gene in the absence of inducer, which can affect normal cell growth) but also for controlling expression of any gene of interest in a straightforward manner, either in its native locus, by replacement of the native promoter, or in a plasmid.[18]

One of the inducible promoters that is more widely used is Pspac. This promoter was developed with the operator and repressor regions of the *lac* operon, and its activation is dependent on the presence of IPTG (isopropyl β-D-1-thiogalactopyranoside). The IPTG inducer binds to the repressor (that was previously attached to the operator region) and modifies its structure, releasing it from the operator region and activating the proximal promoter. Inducible promoters such as Pspac are known to have a certain degree of leakiness, i.e., allowing gene expression even without inducer. Therefore, in the case of inducible/repressor systems, overproduction of the repressor from a high copy number plasmid is often required for proper repression.[18,19]

RNA antisense technology has evolved from the discovery of small interfering RNAs, which regulate several processes in the prokaryotic cell by binding to its sense RNA (target) and interfering with its expression.[20] Generally, antisense RNAs are small, untranslated and highly structured, with one to four stem loops.[20]

Antisense RNAs (asRNAs) can be often designed to repress a gene of interest by binding to its corresponding ribosome binding site and/or start codon region, preventing translation initiation and/or signaling mRNAs for degradation.[21] For successful gene silencing, this binding has to occur before ribosomes' attachment to the mRNA; therefore, optimization of asRNA structure and maximization of its expression is required to increase the molecule stability and distribution in the cell.[21] asRNAs can be either constructed *in vitro* and introduced into the cell (although it is more costly and unpractical for large bacterial cultures) or expressed by the cell using strong promoters or high copy number plasmids.[21] The application of this gene silencing technique can be further developed for determining the essentiality of a gene, by controlling the asRNA expression with inducible promoters. In these cases, the leakiness and inducibility of the promoter should be considered, since it could affect the results.[18] This kind of assays have been useful to assess essential genes in *S. aureus*, as shown in previous studies.[22,23] Although antisense technology has advantages in potentially controlling a certain gene without recurring to chromosome modification, the silencing efficacy varies greatly depending on the targeted gene.[21] Therefore, it is difficult to predict the asRNA fragments that successfully work, and it is often necessary to test several of them before choosing the ideal one.

For assessing the role of a protein, it is usually useful to deplete its expression.[18] Systems of allelic replacement are frequently used methods for depleting/altering the expression of a gene of interest. Vectors designed for that purpose usually carry homologous fragments to the flanking regions of the gene to be deleted, in order to promote single or double crossover recombination events that can either remove the gene while introducing another DNA fragment or simply remove it without inserting other sequences.[18] Different allelic replacement vectors are currently available, that have been developed either as integrative or integration/excision plasmids depending on the purpose of study.

Integrative plasmids usually require a single crossover event in order to be introduced into the chromosome. This kind of vectors are often used to deplete gene expression in one step, as has been described for *Bacillus subtilis* with pMUTIN.[24] This non-replicative vector can also be used in *S. aureus* and carries a *lacZ* gene, to allow the measurement of the expression of the gene to be inactivated, along with an inducible

Pspac promoter, that can be induced by IPTG. By cloning an internal fragment of a gene of interest (or target gene) into pMUTIN and using the resulting plasmid to transform *S. aureus* cells, chromosomal integration of the vector can occur by a single crossover event, inactivating the target gene and fusing it with Pspac as shown in Figure 1. Thus, genes located downstream of the transcriptional fusion that is generated, are controlled by Pspac in an IPTG-dependent manner.[24,25] Integrative plasmids are not recommended for markerless mutation/inactivation of a certain locus (useful for wild-type strains with reporter genes), since its presence in the chromosome frequently implies marker expression.

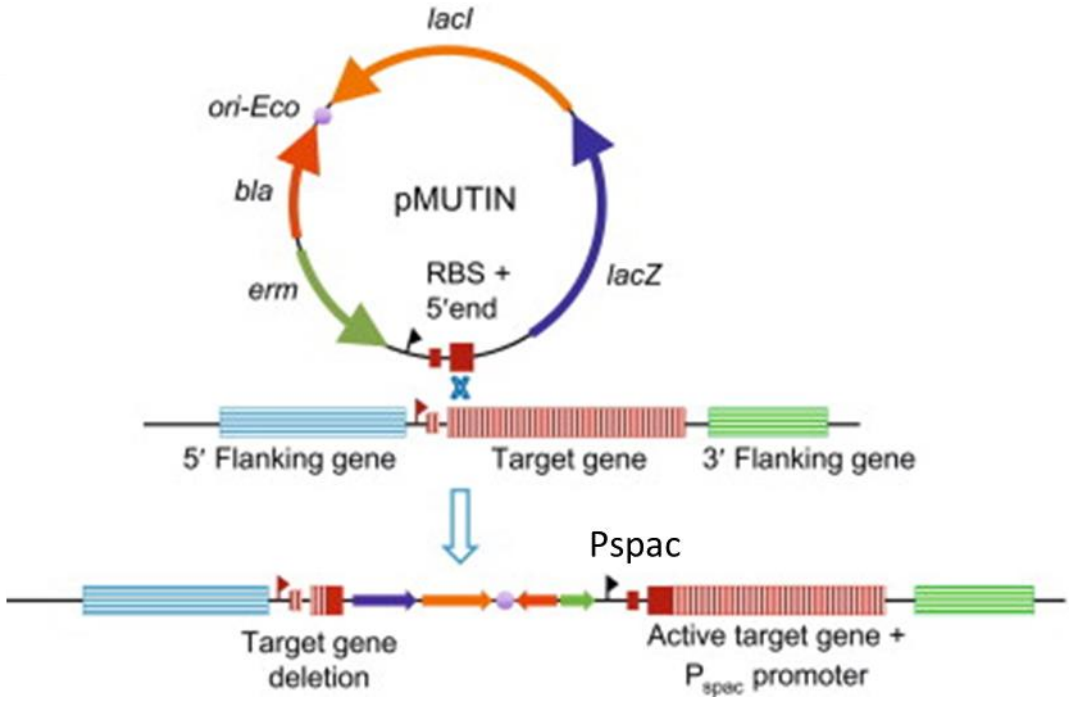



**Figure 1. Integration of pMUTIN through recombination.** Scheme of a single crossover event between a target gene (box in red stripes) and a corresponding homologous 5'end region found in pMUTIN (box in red), leading to the integration of pMUTIN vector into the target chromosome. The resulting integration causes target gene deletion, as well as integration of *lacZ* (β-galactosidase gene), *lacI* (*lac* repressor), *bla* (β-lactamase gene) and *erm* (erythromycin resistance gene), represented by violet, orange, red and green arrows respectively. *ori-Eco* (lilac circle) is the origin of replication for plasmid propagation in *E. coli*. The presence of Pspac promoter (black flag), with a corresponding RBS in the chromosome, after pMUTIN integration, leads to control of target gene expression by addition of IPTG. Adapted from Harwood et al. [26]

Integration/excision plasmids require consecutive single crossovers, or double homologous recombination, for allelic replacement. The most frequently used integration/excision plasmids in *S. aureus* usually carry thermosensitive origins of

replication, thus relying on temperature shifts to perform allelic replacement: from 30ºC, for plasmid propagation in *S. aureus* cells, to 42ºC, which is non-permissive for plasmid replication. Briefly, after introduction and propagation of an integration/excision vector (carrying a mutated loci of the target gene) in the cells, its integration into the chromosome is prompted by growing the same cells in selective conditions at non-permissive temperatures for replication. Cells are then grown at a lower temperature, in the absence of antibiotic selection, to allow for subsequent excision of the plasmid. In these cases, if a second recombination event takes place at the opposite homologous region (not used for integration), loss of the plasmid occurs, keeping the mutated loci in the chromosome.[16] An example of this type of vectors is pMAD, an *E. coli-B. subtilis/S. aureus* shuttle plasmid designed by M. Arnaud and colleagues to perform gene replacement in bacteria as shown in Figure 2.[25] This vector includes a thermosensitive origin of replication that prevents replication when the cells are grown at 42ºC. Additionally, pMAD also encodes a constitutively expressed thermostable ß-galactosidase that allows for colorimetric screening of recombination mutants on X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) plates. White colonies on such plates have undergone successful gene deletion and loss of plasmid, whereas blue/light blue colonies still have presence of the plasmid.[16,18,25] Although this kind of vectors, such as pMAD, are among the most used for locus mutation or knock-out, they still have some disadvantages that must be considered for any potential application. Allelic replacement by integration/excision plasmids is a laborious, time-consuming process that requires several strictly followed steps. It should also be noted that incubations at higher temperatures prevent the construction of thermosensitive mutants.

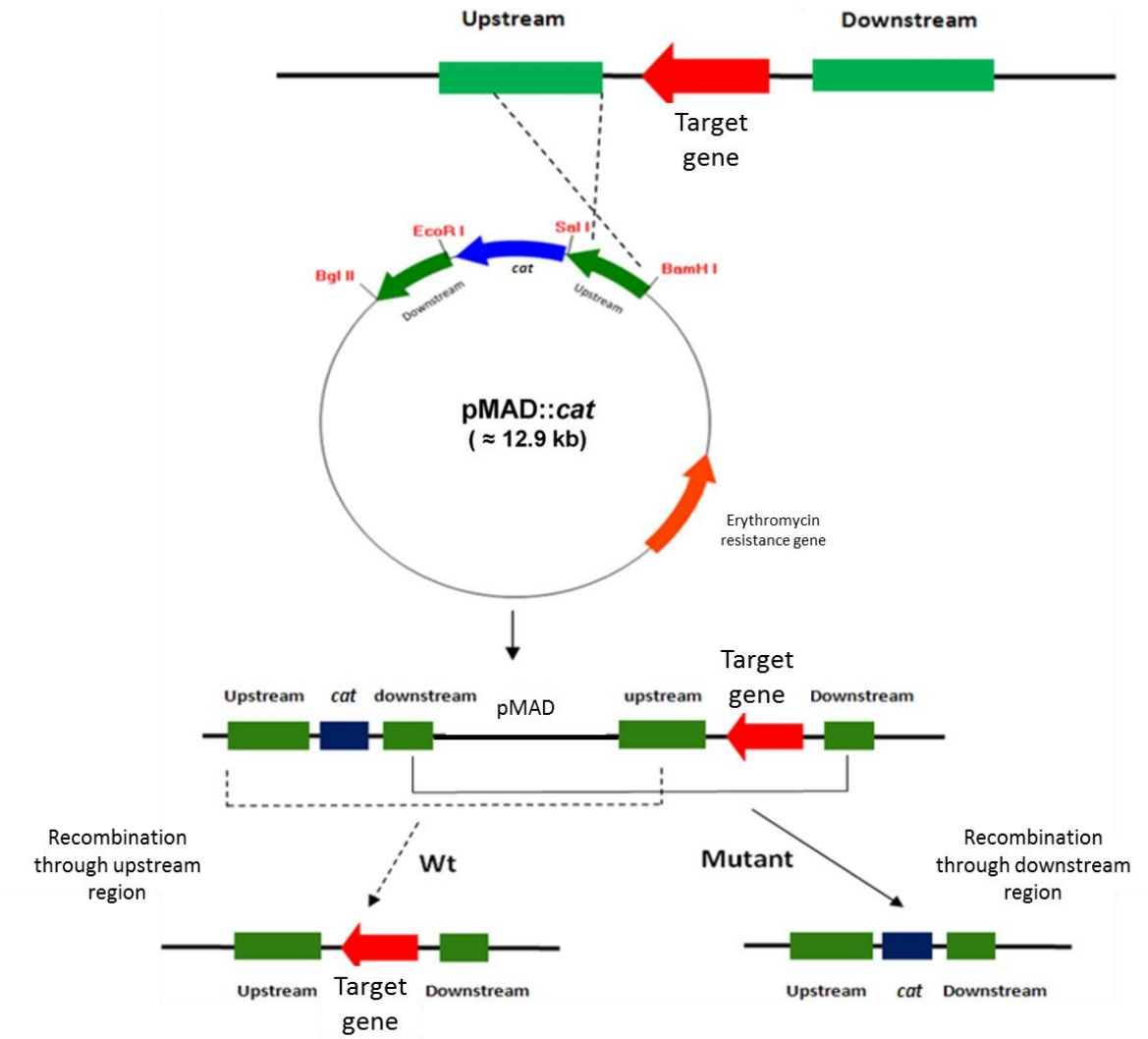



**Figure 2. Allelic replacement of a target gene using pMAD.** Scheme of a pMAD vector encoding *cat* gene (chloramphenicol resistance gene, represented in blue box and arrow) and homologous regions corresponding to the upstream and downstream parts (green boxes) of a target gene (red arrow). Recombination between the upstream regions found in pMAD and the target chromosome leads to the integration of pMAD into the chromosome. If a second crossover event occurs at the same region of the first crossover (in the upstream region), the plasmid is excised and the target gene is restored, originating wild-type (wt) colonies. If the second crossover occurs at the opposite region (downstream), the plasmid with the target gene is excised, leaving *cat* behind. Thus, the resulting colonies do not present the target gene and are resistant to chloramphenicol. Adapted from Bae et al. [27]

**pKILL plasmid**

All of the techniques described above present several pitfalls besides their advantages. Therefore, it is necessary to keep developing genetic tools that can make the study of *S. aureus* easier and less troublesome.

With this purpose, one of the genetic tools we tested in this study, besides CRISPR (explained in further detail in the next chapter), was the pKILL plasmid. This vector was designed and constructed by M. Bramkamp (LMU-Munich) to serve as an allelic replacement system such as the ones we previously described, without resorting to temperatures shifts and laborious steps.

Its structure resembles the allelic replacement plasmids mentioned before, carrying homologous regions flanking a certain gene X of interest (upstream and downstream) and, in between those regions, the DNA fragment with which we want to replace gene X, as seen in Figure 3. The replacement of gene X with the desired DNA fragment occurs through a double crossover event between the homologous regions cloned in pKILL plasmid and the bacterial chromosome, resulting in mutant colonies. This system can be used to make clean knock-out of genes (leaving no resistance marker in the genome) or to replace a gene by a different allele (including variants encoding a fluorescent fusion), or by a resistance marker. The great advantage of this system comparing with other allelic replacement plasmids consists in the fact that integration of a DNA fragment into the chromosome occurs at normal growth conditions (37ºC) in a single step, since pKILL plasmid selects for double crossover mutants by encoding a *Bacillus subtilis* toxin, Yqdb (constitutively expressed by promoter P43).[28] (M. Bramkamp, unpublished) This toxin will be expressed in case a double crossover does not occur, namely if only one single crossover takes place, integrating the whole plasmid in the chromosome and killing the cell, as shown in Figure 3. Furthermore, pKILL plasmid harbors only an *E. coli* origin of replication, making this vector replicative in *E. coli* but not in *S. aureus*. The construction of this plasmid is explained in further detail in Materials and Methods.

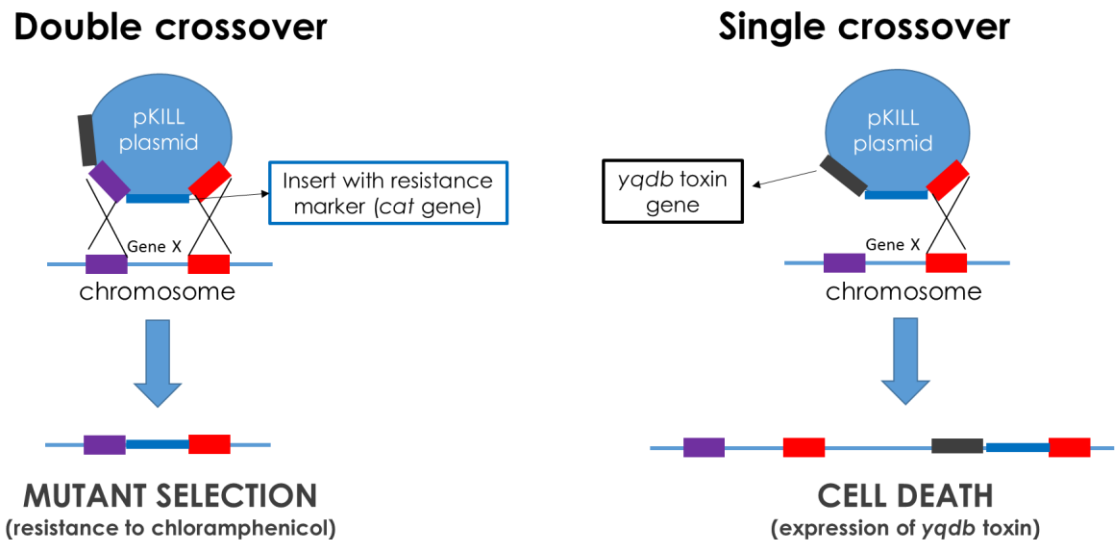



**Figure 3. Representation of the selection process for mutants constructed using pKILL plasmid**. Scheme on the left shows a double crossover event, between homologous regions upstream and downstream of Gene X, found both on the chromosome and cloned in pKILL plasmid. Gene X is then replaced by the *cat* insert found between the homologous regions in pKILL and successful recombinant colonies become resistant to chloramphenicol. Scheme on the right shows a single crossover event. The integration of the whole plasmid in the chromosome allows for *yqdb* expression, killing the cells.

So far, pKILL plasmid has only been tested in *B. subtilis*. (M. Bramkamp, unpublished) As pKILL plasmid is an accessible system to replace genes of interest in a single step, without recurring to temperature changes, it can be a valuable tool for genetic manipulation of *S. aureus*. Therefore, our purpose was to test whether this vector would indeed be useful for *S. aureus*.

## An introduction to CRISPR immunity

CRISPR (clustered, regularly interspaced short palindromic repeats) arrays comprise direct repeats separated by variable spacers that, along with CRISPR associated (cas) proteins, confer resistance against foreign DNA.[29] They are distributed in the genome of several bacteria and almost all archaea.[29]

A CRISPR array was first discovered in 1987 by Ishino and colleagues, who reported 14 repeats that were interspersed by 32-33bp variable spacer sequences located proximal to the isozyme-converting alkaline phosphatase in *E. coli*.[29] However, only 20

years later more progress was made. Two key discoveries - one in 2000 when Mojica and colleagues identified CRISPR arrays in other bacteria and archaea, and the second one in 2005 when several groups observed that many spacer sequences were actually derived from exogenous DNA (such as phage or plasmid DNA) – gave further insight into CRISPR function as a possible adaptive defense system in bacteria.[30] In 2007, this hypothesis was experimentally confirmed by a study which described CRISPR systems as protecting *Streptococcus thermophilus* against lytic phage.[30] Further detailing of this defense mechanism was also published by Marrafini and colleagues, defining DNA as the main target for CRISPR immunity.[31]

Since then, CRISPR arrays have been found in many other bacteria (in 40% of sequenced bacterial genomes and also plasmids) and were established as an immune defense system in prokaryotic cells.[32] CRISPR arrays are now known to typically comprise 23-50bp identical repeats (containing, in some cases, a conserved motif in its 3'end and 5-7bp palindromes, necessary to form stable secondary structures) interspersed with 17-84nt heterogeneous spacers.[29,32,33] Repeats were also shown to be highly conserved within a certain CRISPR locus, but can vary among microbial species.[33] All CRISPR loci present a set of *cas* genes, located less than 1kb away from the array, which encode a heterogeneous family of proteins carrying nuclease, helicase, polymerase and polynucleotide-binding domains.[32,33] In general, bacteria can have more than one CRISPR locus, normally acquired by horizontal transfer by plasmids, megaplasmids and even prophages.[33] A typical CRISPR locus is shown in Figure 4.

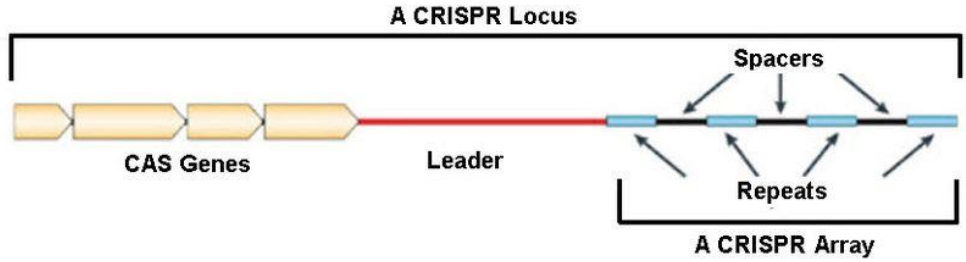



**Figure 4. Composition of a CRISPR locus.** A typical CRISPR locus comprises a CRISPR array, constituted by a series of repeats interspersed with variable spacers. At the 5'end of the array, a leader sequence is found. Cas genes are on the proximal region of the CRISPR array. Adapted from Sorek et al.[29]

The process by which a CRISPR array, along with Cas proteins (CRISPR-Cas), provides protection against invasive elements and serves as a determinant factor for recognizing foreign DNA, was extensively described in several studies.[31–34] This type of systems maintains genetic integrity by allowing the uptake and conservation of advantageous foreign DNA and simultaneously recognizing other invasive DNA which could interfere with the normal cell growth, targeting it for degradation.[33] Briefly, immune defense by CRISPR-Cas systems consists of three different stages: adaptation, expression and interference, shown in Figure 5. The adaptation, or acquisition, phase begins with the exposure to foreign DNA, which is recognized differently from native DNA and cleaved by Cas proteins to promote the insertion of a corresponding short nucleotide sequence as a new spacer into the bacterial host CRISPR array (Figure 5)**.**[30] This way, all spacers in a CRISPR array are homologous to a certain nucleotide sequence (or protospacer) present in an "invader DNA".[35] Acquisition of new spacers does not seem to have a fitness cost for the host cell, probably due to control of the expansion of CRISPR locus. Internal spacer deletions, likely through homologous recombination between CRISPR repeats, were described in the 3' end (where the older spacers can be found) thus possibly allowing for the CRISPR locus to acquire new spacers. New spacers can be derived both from the coding and non-coding strands of the invader protospacer and are added downstream of the leader. The leader, located at the 5'end of the CRISPR array, is an A/T rich, non-coding, sequence which possibly serves as the recognition motif for the addition of new spacers. This sequence is not conserved between species and likely acts as the transcription promoter for the CRISPR array in the expression phase.[29,33]

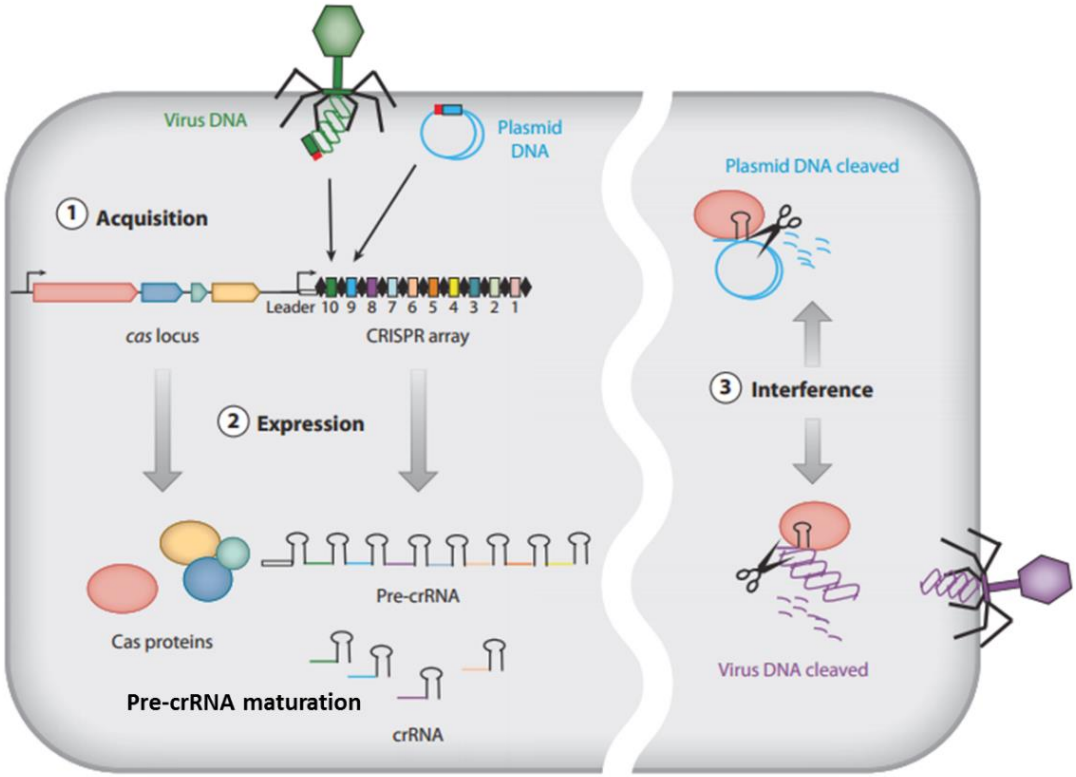



**Figure 5. Stages of CRISPR immunity**. The first stage of CRISPR immunity, Acquisition, consists of the introduction of new spacers – corresponding to sequences found in foreign DNA that enters the cells and is recognized, such as phage or plasmid DNA (green and blue boxes respectively) – into the bacterial CRISPR array. The CRISPR array is located in the chromosome as a succession of spacers (colored boxes) and repeats (black boxes) downstream of a leader sequence, next to *cas* genes. The second phase is Expression of the CRISPR array, originating a pre-crRNA that is processed during pre-crRNA maturation into individual crRNAs, carrying a single spacer. *cas* genes are also expressed at this stage into Cas proteins (colored circles). The final stage, Interference, occurs when foreign DNA (such as plasmid or phage DNA) carrying sequences that are homologous to crRNAs are recognized and cleaved by Cas proteins. Adapted from [36].

The expression phase begins with the transcription of the CRISPR array as a long precursor CRISPR RNA (pre-crRNA), shown in Figure 5. Transcription is constitutive and unidirectional and initiates at the end of the locus containing the leader sequence.[32] In some CRISPR systems, the pre-crRNA is processed into short individual crRNAs (CRISPR RNAs, containing a spacer and flanking repeats) that are bound to a tracrRNA (trans-activating CRISPR RNA).[35] The tracrRNA constitutes a trans-encoded small RNA that contains a partially complementary anti-repeat sequence at the 5'end, which allows the molecule to bind to the repeat region of the crRNA and to form a distinctive secondary structure: upper stem, bulge and lower stem, shown in Figure 6. The 3'end of the

tracrRNA module presents several hairpins required to interact with Cas proteins: the nexus (that often begins with a conserved sequence, UnAnnC, found in several Type II CRISPR-Cas systems) and terminal hairpins (which usually carry a Rho-independent transcriptional terminator), also shown in Figure 6.[30,37] tracrRNAs are thought to co-evolve with Cas proteins and crRNA repeats (shown by tracrRNA precise binding to Cas proteins and species specificity of dual RNA-guided cleavage), and are required for crRNA processing.[35,38] In this process, shown in Figure 7, after the pre-crRNA is transcribed, the complementary portion of the tracrRNA is able to base pair with the repeat portions with the aid of Cas protein. The double stranded RNA complex is afterwards cleaved in the middle of the repeat sequence by native RNase III and other nucleases, forming single repeat-spacer units with a corresponding tracrRNA molecule (crRNA:tracrRNA duplex) and leaving spacer sequence free to base pair with the target DNA in the interference phase.[35,38]

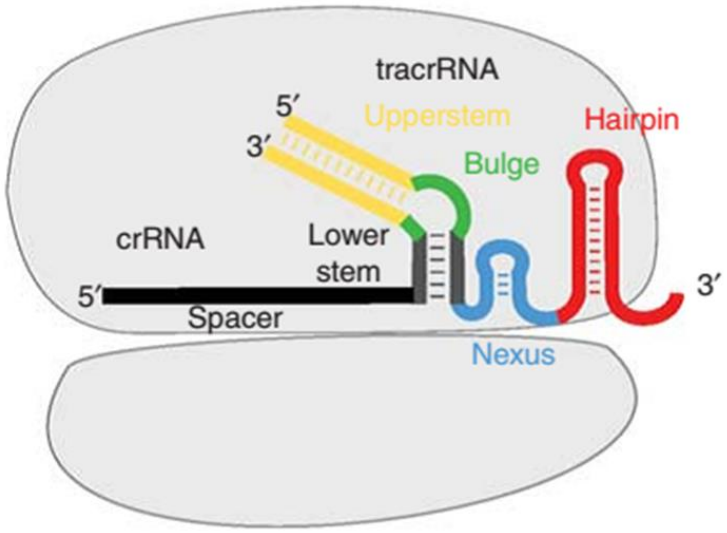



**Figure 6. tracrRNA secondary structure.** Scheme of a crRNA:tracrRNA duplex, inside a Cas protein (in light grey) where the crRNA is represented by black spacer and the tracrRNA is shown with its specific secondary structure: the lower and upper stems and bulge. The nexus and hairpins, necessary for tracrRNA interaction with Cas protein are also indicated in the figure. Adapted from Briner et al.[38]

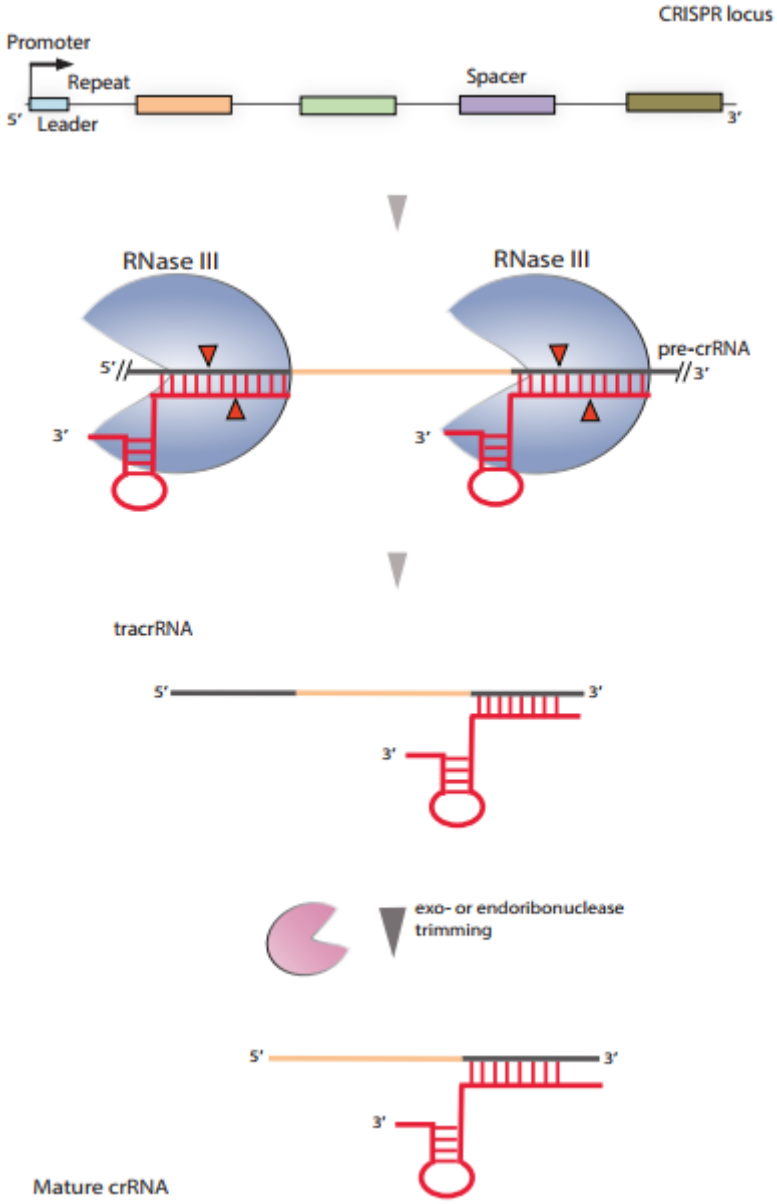



**Figure 7. pre-crRNA processing for Type II CRISPR systems.** Scheme of pre-crRNA maturation into individual crRNAs. A CRISPR locus, or array, is expressed into a pre-crRNA, that carries variable spacers with repeats in between. After the binding of tracrRNA molecules to the repeat regions of pre-crRNA, RNase III trims the pre-crRNA into single crRNAs, which are further trimmed on the 5'end of the spacer by nucleases (in pink), originating mature crRNA with bound tracrRNA. Adapted from Mohanraju et al.[39]

Finally, in the interference phase, the association of crRNAs with Cas proteins takes place to form a ribonucleoprotein complex that will recognize the protospacer (present in the invader's genome and complementary to the crRNA), and cleave it (Figure 5).[30] To avoid auto-immunity, due to the presence of genome sequences similar or

identical to protospacer sequence, recognition of specific short motifs is necessary to restrict the targeting range to only foreign DNA. These sequences, also known as protospacer adjacent motifs (PAM), can be found proximal to the spacer-matching region in the foreign DNA, and are variable among CRISPR systems.[29,35] Thus, they are necessary for DNA recognition and cleavage, and any mutation in its sequence could prevent CRISPR immunity.

A high variety of CRISPR-Cas systems exists amongst diverse bacterial species, differing in protein composition and crRNA processing, but maintaining its basic concept and function. Several classifications have been proposed in order to group the similar CRISPR-Cas systems together, starting with the one Makarova and colleagues proposed in 2011 based on the presence of a signature Cas protein: Type I systems include Cas3, Type II Cas9 and Type III, Cas10. This classification has since been updated and now also considers the composition of effector ribonucleoprotein complexes (crRNP): class 1 systems employ crRNPs with multiple Cas proteins, while class 2 systems include crRNPs with only one Cas protein.[40] Thus, Type I CRISPR-Cas systems belong to class 1; Type II to class 2; and Type III to class 1. Type I systems are characterized by the presence of Cas3, which introduces a single-stranded break (SSB) on the non-target strand of the protospacer. This kind of systems includes a CRISPR-associated complex for antiviral defense (Cascade), a multiprotein crRNA complex that also includes Cas3, and is responsible for detecting and interacting with the PAM on the 5'end flanking region of the protospacer, before cleavage.[40] Specialized Cas proteins are also required in this kind of systems to process the pre-crRNA.[35]

Type II CRISPR-Cas systems, as shown in Figure 8, include a Cas9 protein, which is sufficient for crRNA-guided target DNA cleavage, and that is also involved in the processing of pre-crRNA.[35,40] In the interference phase, Cas9 requires a crRNA:tracrRNA dual complex to identify a PAM on a target DNA and introduce a double-strand break (DSB) into the protospacer if there's complete or sufficient complementarity to the crRNA.[40] Since it does not require any other additional proteins, Cas9 was chosen for a wide use in several applications, as later described.

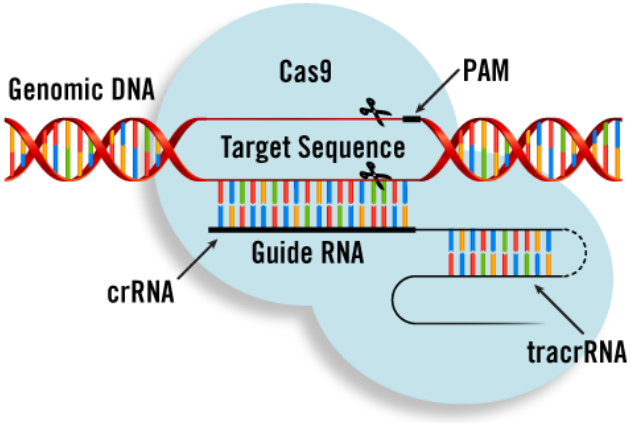



**Figure 8. Representation of a Type II CRISPR-Cas system binding to its target**. Cas9 requires a specific PAM sequence, along with a complementary crRNA:tracrRNA complex, to bind and cleave a certain target. Adapted from [41].

Type III systems are defined by the presence of Cas10 containing PALM domain, which recognizes RNA, thus allowing to target ribonucleotides as well as DNA.[40] This kind of systems does not require PAM for specific cleavage: if the target has complementarity to the repeat regions, it is left intact.[40] Type III systems, like Type I, also use crRNA with multiple Cas proteins as effector complex to recognize and cleave targets.[35]

Recently, other CRISPR-Cas types have been proposed but still need more characterization. Type IV systems are classified into class 1, and are characterized by lacking CRISPR arrays per se. The effector crRNP complex in these systems seems to consist of single copies of Cas5, Cas7 and Csf1 and a putative small subunit.[40] Type V CRISPR-Cas systems belong to class 2 and include, as Type II, a single protein (Cpf1, C2c1 or C2c3, according to the subtypes) that along with crRNA, forms a crRNP complex – without needing the tracrRNA. These proteins also require PAM on the 5'flanking region of the protospacer. Finally, Type VI CRISPR-Cas systems (that belong to class 2) include single-protein effector C2c2 that is guided by crRNAs to the target without tracrRNA. Targets in these systems are RNA sequences, whose cleavage site varies according to its secondary structure.[40]

## CRISPR as a genetic tool

The discovery of CRISPR as a mechanism able to recognize and cleave DNA targets in a specific and straightforward manner was proven to be very useful in genetic engineering. After CRISPR systems' mode of action was established, Jinek and colleagues published an insightful characterization of Type II *Streptococcus pyogenes* Cas9 (SpCas9) showing that this nuclease only requires a crRNA and a tracrRNA for programmable DNA cleavage, thus paving the way for many potential applications.[35,40] The subsequent hybridization of the crRNA-tracrRNA complex into an effective, singly expressed, RNA molecule - the single guide RNA (sgRNA) - allowed the expansion of Cas9 and sgRNA as gene editing tools for prokaryotic and eukaryotic cells.[30,35,40]

After the identification of its catalytically active sites, Cas9 has been modified for different applications, much like sgRNAs, whose stability and spacer sequence proved to be determinant for specific targeting.[40,42] Indeed, soon after the initial genome editing applications of Cas9, off-target effects were taken into account, especially for eukaryotic genomes. Several optimizations, later discussed, were published in order to bypass such concerns.

### sgRNA in CRISPR applications

As described above, sgRNA was designed to hybridize in a single structure the required crRNA-tracrRNA complex for Cas9 guided cleavage**.**[35] Therefore, a sgRNA contains a spacer sequence - also called guide RNA (gRNA), complementary to the targeted DNA sequence - fused to a tracrRNA.[35,43] It was essential to keep the tracrRNA module since it was shown to be necessary for Cas9 functionality: its secondary structure specifically binds to Cas9 (through key nucleotides in the nexus and the lower stem), and restructures the protein into an active conformation**,** enabling Cas9 to interact with the crRNA for a guided DNA cleavage.[35,38] Thus, for the sgRNA construction, the tracrRNA was linked to crRNA by a hairpin structure that retains the base-pairing interactions between both elements, resulting in a final sgRNA molecule that contains the 3'end of the crRNA joined with the 5'end of the tracrRNA (Figure 9). The first applications of this chimeric RNA were mostly in mammalian cells, as shown by several studies.[44,45,46]

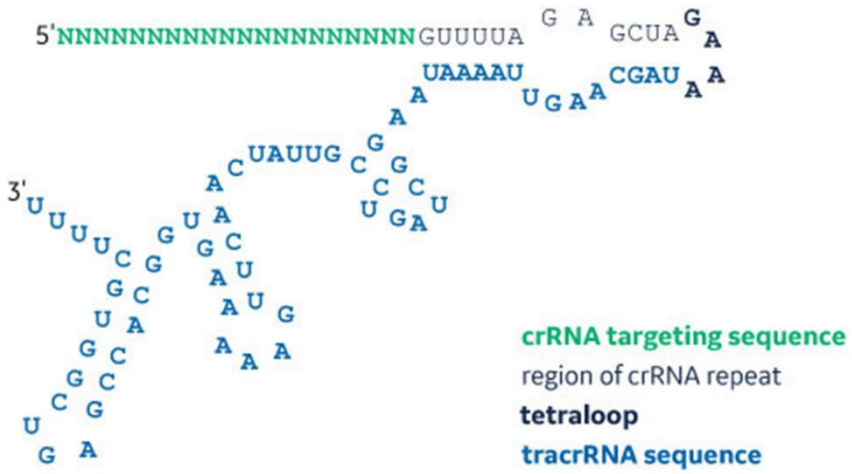



**Figure 9. sgRNA structure.** This RNA molecule comprises a crRNA sequence, which corresponds to the gRNA, fused to a tracrRNA by a tetraloop linker. Adapted from [47].

Given the specificity of the sgRNA for targeting a specific DNA fragment and for interacting with Cas9, it constitutes a determinant factor for CRISPR cleavage success.[42,46] Jinek and colleagues studied how the expression of the sgRNA and/or its loading with Cas9 was the limiting element for Cas9 function in human cells, concluding that higher efficiencies of Cas9 genome targeting could be achieved by optimizing the sgRNA design, expression or localization.[42] According to a later study[48], the sgRNA design has to consider different factors before application: the Cas9 variant to be used, including what PAM sequence it recognizes; the target region or gene that must include the proximal PAM; the promoter to be used for *in vitro* or *in vivo* expression of sgRNA; and the cloning strategy, which might affect the gRNA sequence. Therefore, recent research has focused on sgRNA optimization according to Cas9 proteins and applications.

Two of the first optimizations made were switching a putative U6 Pol-III (RNA polymerase) transcription terminator, TTTT, for TTTA; and extending the Cas9 binding upper stem structure, which enhanced targeting efficiency in some assays with different Cas9 proteins. [37,49,50] Other studies have varied the tracrRNA length (reaching to an optimal 85nt structure), defined the preferred nucleotides in gRNA composition, and also fused proteins with RNA-interacting domains to sgRNA stem loops to increase specificity.[40,43,51] Bioinformatics tools with off-targets scores have also been developed to reduce undesired effects with sgRNA targeting, by predicting potential off-target sequences that could be recognized as well as the targets. These tools are useful for

genome-wide screens that involve thousands of guides or larger genomes, such as in human cells, where off-target effects could be catastrophic. Some of the most widely known prediction tools are bioinformatics applications that search for sequences, with mismatches to the target, that are followed by a PAM motif.[48,50,52]

**Cas9 in CRISPR applications**

Cas9 nucleases, as mentioned before, are a required element for CRISPR applications. There are several variant proteins described, but *Streptococcus pyogenes* Cas9 (SpCas9) remains the most used so far for genome editing. Given that the aim of our study was to develop a CRISPR system for *S. aureus,* we decided to use the *Staphylococcus aureus* Cas9 (SaCas9) which was recently discovered by Ran et al.[53] For summary purposes, only these two proteins will be discussed.

Cas9 proteins constitute a family of enzymes that require a base-paired structure formed between the activating tracrRNA and the targeting crRNA to cleave target dsDNA.[35] Cleavage occurs at locations determined by both base-pairing complementarity between the crRNA and the target protospacer DNA and the presence of a PAM juxtaposed to the complementary region in the target DNA.[35] If the proximal sequence presents enough complementarity to the crRNA – especially in the spacer seed sequence, defined as the minimal required homologous sequence immediately upstream of the PAM – then Cas9 will reach a fully active state for DNA cleavage.[54,55]

Cas9 protein contains two active domains with highly conserved regions between Cas9 variants: RuvC and HNH, each one responsible for the cleavage of one strand within the target DNA.[37,51,56] The bi-lobed structures of SpCas9 and SaCas9 also comprise REC (recognition) and NUC (nuclease) lobes that are converted from the standby to the active state by binding with the sgRNA at the central channel.[51,57,58] Once the Cas9 has paired with the sgRNA, it binds to the target DNA to search for PAM, which is recognized by Cas9 PAM interacting (PI) domain by base pairing of the PAM nucleotides with corresponding aminoacids in the PI. The recognition of the correct PAM leads to interrogation of the proximal DNA strand for sequence complementarity with the gRNA, by base pairing of the gRNA-target DNA duplex in the primary binding channel of Cas9 (between REC and NUC lobes), where complementary target nucleotides interact with

the corresponding aminoacids. In case there is sufficient complementarity, the HNH domain turns about 180º until it is positioned near the third nucleotide position of the target DNA strand proximal to the PAM, while RuvC locates to the same site as HNH but in the opposite strand.[51,59] Once both domains are found in the desired position, they cleave both target and non-target DNA strands, resulting in a DSB (double strand break).[51] This way, Cas9 cleavage requires three stages: the activation step in which Cas9 binds to the sgRNA to search for the target DNA with the corresponding PAM; the repositioning step of HNH and RuvC when the sgRNA-target DNA duplex is formed; and finally the cutting stage, where both strands of the target DNA are cleaved, as shown in Figure 10.[59]

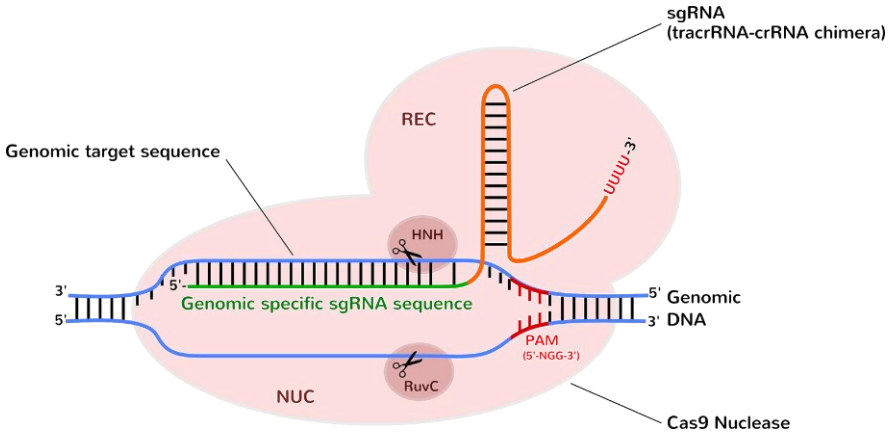



**Figure 10. Cas9-guided DNA cleavage**. A sgRNA, with the gRNA and the corresponding tracrRNA, guides SpCas9 nuclease to a complementary dsDNA target with the specific NGG PAM. Upon recognition and binding, SpCas9 cleaves both strands through HNH and RuvC activity. Adapted from [60].

SpCas9 and SaCas9 recognize different PAMs – thus, they have different PI domains. SpCas9 recognizes an oligonucleotide NGG PAM (that interacts with 5 aminoacids in the PI domain) while SaCas9 recognizes an oligonucleotide NNGRRT PAM (interacting with 7 aminoacids). Interestingly, alteration of aminoacids in the PI domain can change specificity of the Cas9 protein to recognize other PAM sequences, which could be useful for otherwise untargetable sequences.[51,61] Modification of the catalytically active sites in HNH and RuvC domains also alters Cas9 activity. When introducing D10A or H840A mutations (in SpCas9) or D10A and N580A (in SaCas9), these proteins become nickases, which bind and cut only one strand that corresponds to the non-mutated nucleolytic domain.[62] Introduction of both mutations results in a dead Cas9 (dCas9) that still specifically binds to the target DNA but does not cleave it, offering

many potential applications such as *in vivo* imaging and gene expression control. [37,40] Other modifications have also been tested to decrease off target effects, such as altering residues involved in nonsequence-specific contacts with DNA – resulting in high-precision Cas9 proteins - and adding FokI nucleases cleavage domains.[40,50]

SpCas9 was characterized before SaCas9, and became widely used in prokaryotic and eukaryotic organisms, as well as for genome-scale genetic screening.[58] Its oligonucleotide NGG PAM occurs once in every 8bp of random DNA sequence and the subsequent cleavage requires a single-base specificity in 8 to 14bp of the seed sequence, at the 3'end of the gRNA.[55,56,61,63] Up to six contiguous mismatches at the PAM-distal region on the target sequence are tolerated.[35,55] sgRNAs for SpCas9 require a different tracrRNA than SaCas9 and a 20nt gRNA sequence to be effective, although the gRNA can be truncated to 17bp without significantly compromising nuclease activity.[53]

SaCas9 was first described in 2015 by Ran et al., and shortly afterwards applied in several studies, all in eukaryotic cells.[37,53,63] Its main advantage in comparison to SpCas9 (with which it has only 17% nucleotide sequence similarity) is its smaller size (approximately 1kb less than SpCas9), thus allowing its delivery in adeno-associated virus (AAV, cargo size 4.5kb).[53,57] Ran and colleagues assessed SaCas9 cleavage efficiency in comparison to SpCas9, confirming that both were effective at comparable levels, leading to an efficiency of gene modification at over 40%.[49] Subsequent studies also reported high targeting efficiency for SaCas9, even higher in some cases than for SpCas9 depending on gRNA and target sequence.[37,62] Besides an oligonucleotide NNGRRT PAM, specific cleavage by SaCas9 requires a sgRNA with the respective tracrRNA, along with a 21 to 23nt long gRNA (minimal guide length for DNA binding is 11nt) and a 10-12nt seed sequence.[58,63] Single-base mismatches, as for SpCas9, can be tolerated on the PAM-distal region, depending on the identity of mismatched bases.[37] The longer SaCas9 PAM sequence occurs once in every 32bps of random DNA, which limits this protein targeting range but also increases specificity.[63]

Both SaCas9 and SpCas9 share the same REC-NUC bilobed structure with HNH and RuvC domains, previously described.[51,57] Their different PAMs but similar efficiency levels indicate that SpCas9 and SaCas9 can be applied to target distinct DNA sequences in the same cell.[37]

**CRISPR applications**

Applications for sgRNA and Sp/SaCas9 are extremely diversified and are still increasing, as shown by the several published CRISPR-Cas studies so far. Ranging from genome editing in prokaryotic and eukaryotic cells to *in vivo* imaging, its potential seems to be ever-growing.

Genome editing was one of the first uses for CRISPR-Cas systems after Cas9 characterization and sgRNA construction.[64] Essentially, gene editing through CRISPR-Cas systems is achieved by expressing a Cas9 protein, along with a specific sgRNA, to cleave DNA targets that have, at least, 15nt complementarity to the gRNA sequence.[56,59] The expression of this system can be either achieved by introducing an exogenous Cas9 and sgRNA vectors or by recurring to native CRISPR-Cas systems, through the introduction of a pre-crRNA/sgRNA expression cassette.[40] Upon DNA cleavage by Cas9 nuclease or nickase, cells can activate one of two repair pathways to maintain genomic integrity: the non-homologous end joining (NHEJ), in which DNA ends are ligated by ligase-4, or the homologous recombination (HR), where the deleted sequence is restored by copying it from its sister chromatid.[65] NHEJ can further stimulate indel mutations through the addition or removal of a small number of bases by polymerases or exonucleases.[65] Unlike eukaryotic cells, most bacteria lack the NHEJ mechanism (which could rejoin broken fragments resulted from CRISPR interference), and therefore require NHEJ components (such as ligase) and DNA templates, which can be delivered in a plasmid or as linear DNA fragment, to repair DNA ends through homology directed repair (HDR).[40] This way, specific desired mutations can be introduced into the bacterial genome. Furthermore, the double strand breaks (DSBs) induced by Cas9 nuclease, along with HDR, can serve as negative selective markers in bacteria, since DSBs that are not repaired by one of those mechanisms kill the cells – thus leaving only the cells that have gone through successful DNA cleavage and repair, without the need of additional markers in the genome.

For gene editing in eukaryotic and prokaryotic cells, alterations can be applied to sgRNA or Cas9 expression in order to increase specificity or functionality. When multiple loci are to be modified, it could be more convenient to express a pre-crRNA that would be processed (by native or introduced RNase III) into different crRNAs, instead of incorporating several sgRNAs simultaneously.[40] To stimulate recombination events and

decrease cell lethality, Cas9 nickases have been used, as well as recombinase catalytic domains, that were added to Cas9 structure.[40,59,65] Other studies have designed a CRISPR-Cas system that introduces indels without recurring to DSBs, by fusing dCas9 with cytidine deaminase that converts cytosines on the non-target strand to uracils, in mammal cells; and experimented with CRISPR-Cas system AAV delivery in human cells, adding nuclear localization signals to a human codon-optimized Cas9.[40,44,65,66] Finally, self-targeting crRNAs were designed and introduced into bacterial cells to achieve selective killing.[67] These different applications have proven that CRISPR gene editing can be accomplished in different cell lines with relative ease and efficiency; and paved the way for CRISPR therapeutics, which are based in detecting and repairing certain mutations that are associated to human genetic diseases *in vivo*.[59]

For CRISPR therapeutics *in vivo*, it is fundamental to avoid an adverse immune response (treating the patient with immunosuppressive drugs or "humanizing" the Cas9 protein to decrease antibody response), toxicity and also to control off-target activity. Altering an undesired DNA target, for instance, could result in oncogenesis or lead to other unwanted outcomes.[54,59] Additionally, it would be necessary to deal with the ethical responsibility concerns for using CRISPR-Cas9 in human cells, specifically in heritable germline modification.[59] Nevertheless, this possibility opened the door for human clinical trials, starting in 2016 when Chinese doctors removed immune cells from a patient with lung cancer to disable PD-1 gene (which prevents the immune system from killing tumor cells), and then returned the same cells to the body, in order to avoid tumor cells spreading (trial #NCT02793856 in clinicaltrials.gov). This innovative application led to the authorization of a dozen other trials in China with the same purpose – disabling PD-1 gene – in several different cancers. A clinical trial assessing the effects of using a gel containing DNA that encodes for CRISPR machinery against human papillomavirus (HPV) has also been approved (trial #NCT03057912 in clinicaltrials.gov). Furthermore, in a study applied to virulent *S. aureus* strains, delivery of Cas9 and crRNAs targeting *aph-3* kanamycin resistance gene through bacteriophages also implied other potential uses besides viral and oncology treatments.[59]

CRISPR-Cas systems have also been used to control gene expression, through transcriptional repression or activation. Transcriptional repression or gene silencing through CRISPR (CRISPR interference, or CRISPRi) uses dCas9 specifically guided to a target locus to block its transcription, by preventing the movement or the binding of the RNA polymerase (Figure 11).[56,68] Blockage of RNA polymerase binding to the promoter by dCas9, proved to be more effective than targeting the open reading frame, independently of the target strands. In fact, the further the binding site is from the ribosome binding site, the smaller the change in gene expression.[40] Besides dCas9, truncated gRNAs can also be used for CRISPRi. They were shown to guide Cas9 to target sites but not to activate its nuclease activity, thus allowing for alternation between repression and cleavage modes by exchanging full-length and truncated sgRNAs.[40]

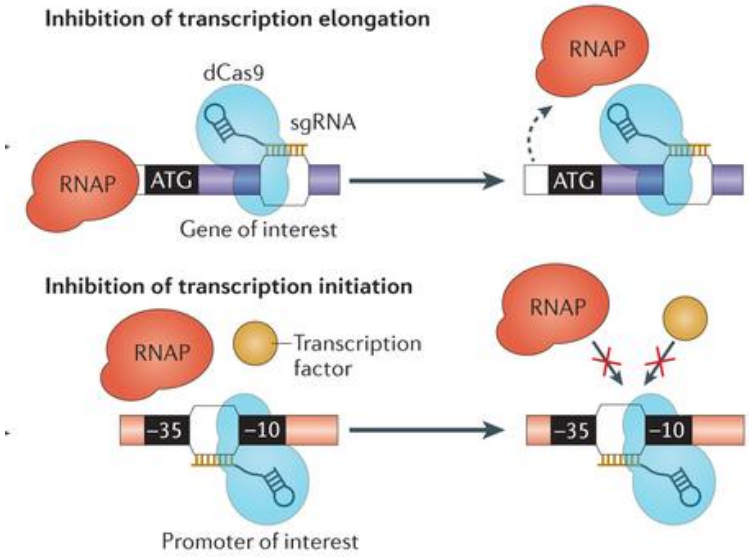



**Figure 11. Mechanisms of dCas9 repression.** dCas9 can bind either to the open reading frame of the gene of interest – blocking RNA polymerase movement – or to the gene promoter – inhibiting RNA polymerase binding. Adapted from Qi et al. [69]

Gene activation through CRISPR can be achieved by fusing dCas9 to certain subunits/domains known to stabilize RNA polymerase binding and stimulate transcription. Bikard and colleagues first reported this application.[53] In that study, the *E. coli* RpoZ subunit (a RNA polymerase subunit involved in its structural stability) was fused to dCas9, increasing up to 23-fold the expression rate of *gfp-mut2* gene when targeted to that reporter gene.[65,70] Interestingly, in this system the rate of transcription activation depended on the native strength of the promoter of the target gene: for weaker promoters, the increase was more dramatic than for strong promoters.[56] Other factors,

such as the target position of the activating dCas9 (400-50bp upstream of the transcription start site is ideal) and the fusion configuration between dCas9 and RpoZ subunit, were also suggested to affect the expression level.[48,56,65] In eukaryotic cells, transcription enhancement was achieved by overexpressing dCas9 fused to transactivation domains like VP16 and p65, activating both reporter and endogenous genes.[71–73]

dCas9 specific binding to the target has also been used for chromosome locus imaging *in vivo*. To tag a genomic locus for live-cell high resolution imaging, it's necessary to design a dCas9:sgRNA complex, containing a fluorescent protein domain, complementary to the desired target.[37] A certain fluorescent marker could be either fused to dCas9 or be directed to specific regions of dCas9 or sgRNA through protein-protein or RNA-protein interaction domains.[48] In the initial studies and since only one GFP (green fluorescent protein) molecule was recruited per dCas9:sgRNA complex, the labeled targets were mostly repetitive elements such as telomeres and *mucin1* and *mucin4* genes, in human cells.[74] To overcome this limitation, several studies in eukaryotic and prokaryotic cells suggested different approaches: tiling 26 to 36 unique sgRNAs throughout non-repetitive loci; adding a super-folder GFP tandem array to dCas9 for stronger fluorescence signal; constructing extended sgRNAs that would recruit several GFP molecules; and designing dCas9 with peptide arrays to recruit antibody fusion proteins (SunTag). [73,75,76] In parallel, the possibility of using different dCas9 orthologs (with different PAM recognition) fused to red, green and blue fluorescent proteins was analyzed in a study by Ma and colleagues, which showed that multicolor imaging in live cells is possible using CRISPR.[77]

Despite all the promising applications that have been developed for CRISPR, off-target effects, besides delivery methods, are currently one of the main concerns for CRISPR gene manipulation, especially in eukaryotes. Different approaches have been suggested in order to reduce off-target effects, such as: (i) use Cas9 proteins from species with larger genomes and that have undergone frequent horizontal gene transfer along with their CRISPR loci, as they are probably the most specific of all; (ii) modify PAM sequences for longer and less occurring motifs; (iii) design sgRNA structures for higher efficiency; (iv) reduce the basal Cas9 affinity for DNA; (v) use paired Cas9 nickases (with corresponding paired sgRNAs).[49,54,64,72] Limiting the duration of Cas9 expression, as well as controlling its delivery modality and dosage, have also been recommended as potential

options for increasing specificity, especially for human cells.[54] Finally, several protocols for the detection of genome-wide Cas off-target activity in cells have been described as important complements to other strategies. In an article written by Tycko and colleagues[54], the application of biased (*in silico* prediction) and unbiased (direct detection of off-targets) methods was suggested, in order to analyze off-target sites *in vitro* or *in vivo* prior to validation *in vivo*. This study also recommends using bioinformatics tools to generate a ranking for sgRNAs with the best genomic specificity.

All the mechanisms that we have here summarized present several advantages in comparison with current gene editing/modulation techniques. Many CRISPR applications are still in early stages but are being developed at an exponentially increased rate, and show promising results that could transform our capacity to manipulate genomes.

# 2. Materials and methods

## Bacterial strains and growth conditions

Strains, plasmids and primers used in this study are described in Table 2, Table 3 and Table 4 respectively, shown in the end of this chapter. *S. aureus* cells were grown at 37ºC with aeration in liquid (TSB; Tryptic Soy Broth, from Difco) or on solid medium (TSA; Tryptic Soy Agar, from VWR), supplemented with antibiotics when necessary: chloramphenicol 10 µg/ml (Cm, from Sigma-Aldrich), erythromycin 10 µg/ml (Ery, from Apollo Scientific) or kanamycin 50 µg/ml (Kan, from Apollo Scientific) and neomycin 50 µg/ml (Neo, from Sigma-Aldrich). *E. coli* cells were also grown at 37ºC with aeration in either LB (Luria-Bertani broth, from VWR) or LA (Luria Agar, from VWR), supplemented with ampicillin 100 µg/ml (Amp, from Apollo Scientific) when required. Strain growth was observed by optical density (OD) measurement at 600 nm.

## Molecular cloning

### DNA purification and manipulation

Total DNA was extracted from *S. aureus* strains after overnight growth at 37ºC on TSA plates. Cells were scraped from the plate, re-suspended in 50 mM Ethylenediaminetetraacetic acid (EDTA, Sigma-Aldrich) with Lysostaphin 10 µg/ml and RNase 20 µg/ml (both from Sigma-Aldrich) and incubated for 30 minutes at 37ºC. Nuclei Lysis Solution (Promega) was added to the lysed cells, which were then incubated for 5 minutes at 80ºC. After cooling the samples to room temperature, Protein Precipitation Solution (Promega) was added. DNA precipitation was then achieved by addition of two thirds of sample volume of isopropanol. The DNA pellet was washed with 70% ethanol and finally re-suspended in water.

Plasmid DNA was extracted from *E. coli* cells using Wizard SV Plus Miniprep Kit (Promega), following the manufacturer's instructions. DNA was digested with FastDigest restriction enzymes (ThermoFisher) by incubating 0.5-1 µg DNA at 37ºC for 1.5 hours, with 1X FastDigest buffer, 1 µl of the selected endonuclease and 1 µl of Fast Alkaline Phosphatase (ThermoFisher), when necessary. DNA fragments were purified using Wizard SV Plus Cleanup Kit (Promega), according to manufacturer's instructions.

When higher concentrations of plasmid DNA were required, plasmid Midipreps were prepared using HiSpeed Plasmid Midi Kit (Qiagen), following the manufacturer's instructions. DNA was then precipitated by adding to the samples one tenth of plasmid DNA's volume of sodium acetate (3 M, pH 5.2), and then 2-3X final volume of ethanol 95%. After an incubation on ice for 15 minutes, the samples were centrifuged for 30 minutes at 20000 x g and the pellet was washed with 70% ethanol. Finally, the DNA pellet was dissolved in sterile water.

DNA ligations were performed by standard molecular biology techniques using T4 DNA Ligase (ThermoFisher) in recommended reaction conditions. PCR reactions were performed by using either Phusion DNA Polymerase for cloning purposes or DreamTaq Polymerase for colony screenings (both from ThermoFisher), following the manufacturer's instructions.

### *E. coli* transformation

*E. coli* DC10B competent cells were prepared according to the Rubidium Chloride ($RbCl_2$) protocol.[79] For that, early exponential cells ($OD_{600nm}$ 0.4-0.5) were incubated on ice for 15 minutes and later centrifuged (7197 x g for 15 minutes at 4ºC). The pellet was re-suspended in 1/3 of culture volume of RF1 buffer ($RbCl_2$ 100 mM, Potassium acetate 35 mM pH 7.5, $MnCl_2$ tetrahydrate 50 mM, $CaCl_2$ bihydrate 10 mM and Glycerol 15%) and incubated on ice for 15 minutes. After ice incubation, the samples were centrifuged as before, and re-suspended in ½ volume of ice-cold RF2 buffer (MOPS 10 mM, $RbCl_2$ 10 mM, $CaCl_2$ bihydrate 75 mM and glycerol 15%). Glycerol was added and the cells were stored at -80ºC in 100 µl aliquots. For transformation, 5 µl of ligation DNA or 1 µl of extracted plasmid DNA were added to thawed competent cells; which were then incubated on ice for 5 minutes, heat shocked for 1 minute at 42ºC, and rescued in 1 ml of LB. Transformed cells were incubated at 37ºC for 1 hour and plated on LA with Amp, growing at 37ºC overnight.

### *S. aureus* transformation

RN4220 electrocompetent cells were prepared as previously described.[80] Briefly, cells were grown in TSB at 37ºC until an $OD_{600nm}$ 0.4 and then harvested by centrifugation (7197 x g for 15 minutes at 4ºC). The obtained pellet was washed with an equal volume

of 0.5 M sterile sucrose, centrifuged, and washed again in ½ volume of sucrose 0.5 M. After this washing step, cells were incubated on ice for 15 minutes, re-suspended in 1/100 of the initial volume of sucrose 0.5 M, divided in 50 µl aliquots and stored at -80ºC. To transform the RN4220 competent cells, the same were mixed with 0.5-17 µg of purified DNA, followed by a 10 minute incubation on ice. The cells were then electroporated (2.5 kV; 25 µF and 100 Ω) using a 0.2 cm BioRad Gene Pulser cuvette, and recovered in 1 ml TSB for 1 hour at 37ºC. After recovering, the cells were plated on TSA with antibiotic.

The transformation efficiency of the competent cells was calculated after electroporation with a high-copy plasmid of known concentration. After the electroporation, all the colonies in one plate were counted, and the final number of transformants per ng of DNA was calculated.

***S. aureus* transduction**

Transductions were performed using phage 80α as previously described.[81] Phage lysates were first prepared by re-suspending the donor strain in 1 ml TSB with $CaCl_2$ 5 mM. Phage 80α was serially diluted to $10^{-7}$ in Phage Buffer ($MgSO_4$ 1 mM, $CaCl_2$ 4 mM, Tris-HCl 50 mM pH 7.8, NaCl 5.9 g/L, gelatin 1 g/L). 10 µl of each phage dilution was mixed along with 10 µl of re-suspended cells in 3 ml of Phage Top Agar (casamino acids 3 g/L, Difco; yeast extract 3 g/L, Difco; NaCl 5.9 g/L, Panreac; agar 5 g/L, VWR; pH 7.8, $CaCl_2$ 5 mM). The mixtures were then poured onto Phage Bottom Agar (same composition as for Phage Top Agar, but with 15 g/L agar instead) with $CaCl_2$ 5 mM, and incubated overnight at 30ºC. Plates that showed almost-confluent lysis were selected and incubated with 3 ml ice-cold Phage Buffer for 1 hour at 4ºC. The top agar, along with the phage buffer, was collected, vortexed to disrupt the agar and incubated for 1 hour at 4ºC. The phage lysate was recovered after centrifugation, filter-sterilized using a 0.45 µm sterile filter and plated (50 µl) on TSA to confirm sterility.

For the transduction, the recipient strain was grown overnight at 37ºC on TSA and re-suspended in 1 ml TSB with $CaCl_2$ 5 mM. Phage lysate was added in different volumes (0.1 µl, 1 µl, 10 µl and 100 µl) to 100 µl of the previously prepared cell resuspension and 100 µl phage buffer. The transduction mixtures were incubated at 37ºC for 20 minutes, added to 3 ml 0.3GL Top Agar (casamino acids 3 g/L, Difco; yeast extract 3 g/L, Difco; NaCl, 5.9 g/L Panreac; sodium lactate 60% syrup, 3.3 ml/L, Sigma-Aldrich; glycerol

50%, 2 ml/L, Sigma-Aldrich; Tri-sodium citrate, 0.5 g/L, Sigma-Aldrich; and agar 7.5 g/L, VWR; pH 7.8) and then poured onto double-layer plates containing on the bottom 10 ml of 0.3GL bottom agar (identical composition to 0.3GL Top Agar, with 15 g/L agar instead) supplemented with 3X the working antibiotic concentration, and on the top 20 ml of 0.3GL bottom agar, without antibiotic. For transductions performed with two antibiotics simultaneously, such as Kan and Cm, a 150 µg/ml concentration for Kan was used, instead of using Kan 50 µg/ml and Neo 50 µg/ml.

## Fluorescence microscopy

Strains were grown overnight at 37ºC in TSB supplemented with antibiotic when required. The following day, the cultures were diluted 1:200 in 10 ml TSB, with antibiotic and inducer when necessary, and incubated again at 37ºC until mid-exponential phase ($OD_{600nm}$ 0.4). 1 ml of culture was pelleted (20000 x g for 1 minute at room temperature) and re-suspended in 20 µl of 1X M9 medium, which was diluted from 4X M9 medium (KH2PO4 13.6 g/L, K2HPO4 11.6 g/L, di-Ammonium citrate 2.8 g/L, Sodium Acetate 1.04 g/L (all from Sigma-Aldrich), Glucose 40 g/L, Casaminoacids 40 g/L (VWR), MgSO4.7H2O 2.8 g/L, CaCl2 28 g/L (Sigma-Aldrich), 4X 50X MEM Aminoacids and 4X 100X MEM Vitamins, both from Thermo-Scientific). 1-2 µl of culture was then placed on a thin layer of 1.2% agarose in 2.5X M9 medium, also diluted from 4X M9 medium. Fluorescence microscopy was performed using a Zeiss Axio Observer.Z1 microscope equipped with a Photometrics CoolSNAP HQ2 camera (Roper Scientific), using phase contrast objective Plan Apo 100 x/1.4 oil Ph3, with 0.24 µm resolution and 0.55 numerical aperture. The software used was either Metamorph (Molecular devices) or Zen blue software (Zeiss). Analysis of fluorescence images was performed using Metamorph and ImageJ softwares.

## Testing the pKILL plasmid

### Construction of pKILL plasmids

The pKILL plasmids that we have tested in *S. aureus* were constructed and kindly provided by M. Bramkamp (LMU-Munich). Their relevant characteristics are summarized in Table 3. For their construction, a BsaI restriction site was eliminated from

pUC18, and the *yqdb* toxin gene (under the control of the constitutive promoter P43) was cloned between restriction sites KpnI and BamHI. During that same cloning step, a BsaI multiple cloning site was also introduced to insert the corresponding elements of each tested pKILL plasmid (pKILL TatA and pKILL FloA) via golden gate cloning. All of pKILL plasmids have been sequenced to confirm the presence of its components.

**Introduction of pKILL FloA and pKILL TatA in *S. aureus***

pKILL FloA and pKILL TatA were tested by electroporation of RN4220 competent cells following the protocol described above. pKILL FloA test electroporations A and B were performed using 5.6 and 6.0 µg of DNA, whereas pKILL TatA electroporations A, B and C were performed with 6.8, 13 and 17 µg DNA respectively (see Results and Discussion). As positive control, RN4220 competent cells were transformed with 0.12 µg of the high-copy plasmid pEPSA5.[82] Transformed RN4220 cells were plated on TSA with 20 µg/ml Cm and incubated at 37ºC for up to two overnights, in order to guarantee mutant colonies' growth. Grown colonies were afterwards streaked on MSA (Mannitol Salt agar, from Difco) and new TSA plates with Cm (10 µg/ml). The pKILL colonies that grew on TSA plates with Cm, and on MSA with a distinctive yellow halo were screened through phenotypic and genotypic analysis.

For the phenotypic analysis of RN4220 pKILL TatA, colonies were streaked on MSA, TSA with Cm and TSA plates supplemented with egg yolk, which were prepared by adding half of an egg yolk to 200 ml melted TSA. The cells that grew on the described plates and presented alterations in their degradation halo on egg yolk agar, were then confirmed by genotypic analysis. For RN4220 pKILL FloA colonies, phenotypic analysis was performed by streaking the cells in MSA and TSA with Cm. The colonies that grew on such plates were grown overnight in TSB with Cm (10 µg/ml) and observed the following day, after inoculation in 10 ml TSB, by fluorescence microscopy as described above. The RN4220 pKILL FloA colonies with a FloA-mNeongreen signal were submitted to a genotypic analysis.

Genotypic analysis of RN4220 pKILL TatA and RN4220 pKILL FloA was performed by colony screening PCR to confirm allelic replacement (see Table 1 for PCR conditions). Primers used were FloA and TatA sequencing primers (P1/P2 for TatA and P3/P4 for FloA), complementary to up- and downstream regions of *tatA* and *floA* genes respectively.

In cases where colony PCR did not show the expected knock-in band for pKILL FloA, primers testing the incorporation of the full plasmid in the genome were used: P5/P6 (amplifying a 1 kb fragment from ampicillin resistance gene) and P7/P3 or P8/P4, where P7 and P8 are complementary to *cat* gene. P9 and P10 primers, specifically designed to confirm a *S. aureus* strain, were also used in case of inconclusive results using the MSA plates. Each primer is described in Table 4, shown at the end of this chapter.

**Table 1. PCR conditions for colony screening of RN4220 pKILL TatA and RN4220 pKILL FloA**

| Step | Temperature | Duration | Number of cycles |
|---|---|---|---|
| **Initial denaturation** | 95ºC | 5 minutes | 1 |
| **Denaturation** | 95ºC | 30 seconds | 30 |
| **Annealing** | 58ºC | 30 seconds | |
| **Extension** | 72ºC | 4min30sec (pKILL FloA); 2min45sec (pKILL TatA) | |
| **Final extension** | 72ºC | 7 minutes | 1 |

## Designing the CRISPR system

### Construction of *S. aureus* strains expressing dCas9-sfGFP

D10A and N580A mutations in *S. aureus* Cas9 protein sequence were previously reported to originate dCas9, a non-functional nuclease.[37] To obtain *dCas9* gene sequence, the same mutations were introduced into *S. aureus cas9* gene sequence (kindly provided by F. Ann Ran (MIT)) by replacing an adenine for a cytosine in codon 10 and two adenines for a guanine and cytosine in codon 580 respectively. To construct the dCas9-sfGFP fusion, we designed a DNA sequence that was synthetized by NZYTech (Portugal) comprising: a FtsZ-derived ribosome binding site, a start codon, the *dCas9* gene sequence, and a stop codon. Afterwards, this DNA fragment was digested with SalI and BamHI restriction enzymes and cloned into pCNX-GFP plasmid, giving rise to pCNX-dCas9-GFP. pCNX-GFP plasmid was previously constructed by amplifying the *sfgfp* gene sequence from pTRC99a-P7 plasmid[83] with Phusion polymerase, using primers P11 and P12 (Table 4), under the following PCR conditions: an initial denaturation step of 98ºC

for 30 sec followed by 25 cycles of amplification, composed of denaturation at 98ºC for 10 sec, annealing at 56ºC for 30 sec, extension at 72ºC for 30 sec and a final extension step at 72ºC for 7 min. Afterwards, this PCR product was digested with BamHI and EcoRI and cloned into pCNX[84], and confirmed by sequencing.

The dCas9-sfGFP fragment cloned in pCNX-dCas9-GFP was sequenced and this plasmid was electroporated into RN4220 (selection with 50 µg/ml Kan/Neo) and transduced to NCTC8325-4 FP650-RodZ and COL MurJ-mCherry, resulting in NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP and COL MurJ-mCherry pCNX-dCas9-GFP.

**Construction of plasmids for expression of sgRNA targeting FP650-RodZ and MurJ-mCherry**

The sgRNA1 and 2 designed to target the non-ATG and the ATG strands of *fp650-rodZ* respectively were expressed from pGC2[85] vector under the control of *pbpb* (PBP2) promoter. To construct pGC2-sgRNA1, we amplified the *pbpb* promoter sequence (region corresponding to P1 as described by Pinho et al.)[86] from previously extracted NCTC8325-4 genomic DNA by using P13 and P14 primers (Figure 12). P14 includes the sgRNA1 sequence, which besides the gRNA includes the tracrRNA sequence for *S. aureus* Cas9, with a terminator, as found on BPK2660 (Addgene plasmid #70709). PCR with these primers and Phusion polymerase was performed under the same conditions as for *sfgfp* gene, to obtain a DNA fragment containing *pbpb* promoter and sgRNA1 (Figure 12). The resulting PCR product was digested with SmaI and EcoRI, cloned into pGC2 (Table 3, at the end of this chapter), and confirmed by sequencing.

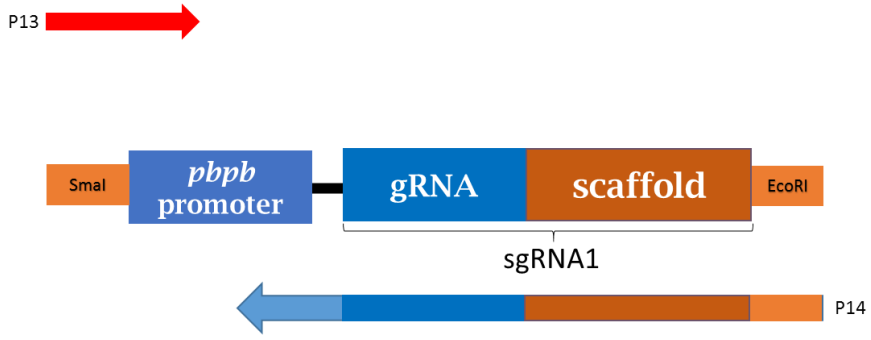



**Figure 12. Scheme of sgRNA1 amplification by PCR.** PCR was performed to obtain the final insert, SmaI – *pbpb* promoter – sgRNA1 – EcoRI, here represented in the scheme. We amplified *pbpb* promoter from NCTC8325-4 genome using primers P13/P14 (Table 4). P14 is complementary to the 3'end of *pbpb* promoter (in light blue) and includes the sgRNA1 sequence, which comprises the gRNA (dark blue) and

scaffold (dark orange) sequences. The scaffold includes the tracrRNA and terminator sequences from BPK2660.

To obtain a pGC2-sgRNA2 vector under the control of *pbpb* promoter, we amplified two DNA fragments from NCTC8325-4 genomic DNA and pGC2-sgRNA1 using primers P13/P17 and P16/P18, respectively (Figure 13) using the same PCR conditions as above. Afterwards, these fragments were joined by overlap PCR to obtain a single DNA fragment comprising *pbpb* promoter and sgRNA2, using primers P13/P18 (Figure 13). The resulting PCR product was digested with SmaI and EcoRI, cloned into pGC2 (Table 2, at the end of this chapter), and confirmed by sequencing.

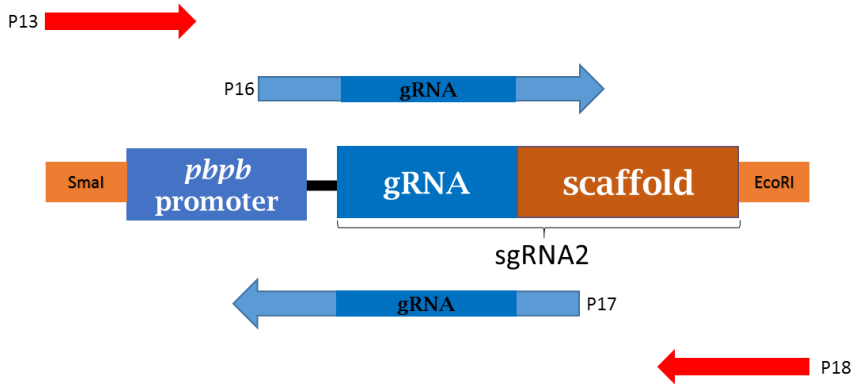



**Figure 13. Scheme of sgRNA2 amplification by PCR.** An overlap PCR using primers P13/P18 was performed to obtain the insert, SmaI – *pbpb* promoter – sgRNA2 – EcoRI, here represented in the scheme. For that, first the promoter was amplified from NCTC8325-4 genome using primers P13/P17, and sgRNA2 was amplified from pGC2-sgRNA1 using primers P16/P18 (Table 4). Primers P17 and P16 include the new gRNA sequence for sgRNA2.

To control the expression of sgRNA3 and sgRNA4 (targeting the promoter and 3' end regions of the ATG strand of *fp650-rodZ*, respectively) and MurJ sgRNA (targeting 5' end of the ATG strand of *murJ-mCherry* gene), a synthetic minimal promoter, J23119, was chosen (BBa_J23119).[68] pGC2 plasmids expressing each of these three sgRNAs were obtained by first amplifying the scaffold sequence from pGC2-sgRNA1 plasmid with primers P19/P18, P20/P18 and P21/P18, where P19, P20 and P21 include the J23119 promoter together with sgRNA3, sgRNA4 or MurJ sgRNA respectively (Table 4, Figure 14). PCR conditions were identical to those for sgRNA2. Each DNA product was then digested with SmaI and EcoRI, cloned into pGC2 (Table 2, at the end of this chapter), and finally confirmed by sequencing.

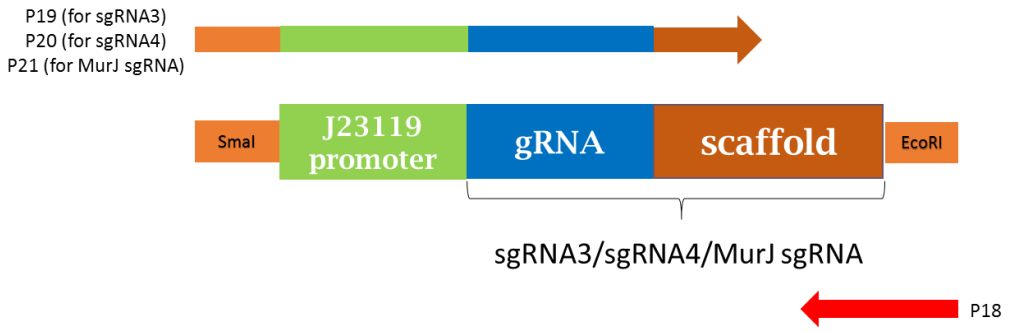



**Figure 14. Scheme of sgRNA3/sgRNA4/MurJ sgRNA amplification by PCR.** PCR was performed to obtain the final insert, SmaI – J23119 – sgRNA3/sgRNA4/MurJ sgRNA – EcoRI, here represented in the scheme. The insert was amplified from pGC2-sgRNA1 using primers P19/P18 for sgRNA3; P20/P18 for sgRNA4 and P21/P18 for MurJ sgRNA (Table 4). P18 is complementary to the 3'end of the scaffold, whereas P19/P20/P21 are complementary to the 5'end (dark orange) and include each the corresponding sgRNA ( dark blue) fused to J23119 promoter (green).

Each sgRNA plasmid for FP650-RodZ (sgRNA1 to sgRNA4) and for MurJ-mCherry (MurJ sgRNA) was electroporated into RN4220, selected on Cm plates, and transduced into the desired *S. aureus* cells (see Construction and microscopy observation of CRISPR strains).

**Construction of NCTC8325-4 FP650-RodZ and COL MurJ-mCherry**

COL MurJ-mCherry (V. Macário and M. G. Pinho, unpublished; Table 2) was constructed by first amplifying three DNA fragments (F1, F2 and F3) from COL (primers P22/P23 for F1, P24/P25 for F2 and P26/P27 for F3, Table 4) that were joined by overlap PCR, giving F1-F2-F3, corresponding to the 3' end region of MurJ, without STOP codon, fused to mCherry. This PCR product was digested with BgIII and BamHI and cloned into pMAD giving rise to pMAD-MurJmCherry, which was electroporated into RN4220 at 30ºC (selected with erythromycin) and subsequently transduced to COL (selection at 30°C with erythromycin). The replacement of *murJ* gene for *murJ-mCherry* was completed after a two-step homologous recombination as previously described.[25]

For the construction of NCTC8325-4 FP650-RodZ (H. Veiga and M. G. Pinho, unpublished; Table 2), three DNA fragments (F1, F2 and F3) with overhangs complementary to the adjacent fragments, were amplified by PCR, using primers P28/P29 for F1, P30/P31 for F2 and P32/P33 for F3 (Table 4). These three fragments were joined together by overlap PCR, resulting in F1-F2-F3, that comprises eqFP650 (without STOP

codon) fused to *rodZ* gene by a linker. This PCR product was then digested with SalI and NcoI and cloned into pMAD, originating pMAD-FP650RodZ that was sequenced. pMAD-FP650RodZ plasmid was electroporated into RN4220 at 30ºC (selected with erythromycin) and subsequently transducted to NCTC8325-4 (selection at 30°C with erythromycin). The replacement of *rodZ* gene for *eqfp650-rodZ* was completed after a two-step homologous recombination, as described before.[25]

**Construction and microscopy observation of CRISPR strains**

NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP and COL MurJ-mCherry pCNX-dCas9-GFP strains received by transduction the vectors expressing the sgRNAs. Briefly, NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP was transduced with (i) pGC2-sgRNA1, (ii) pGC2-sgRNA2, (iii) pGC2-sgRNA3 and (iv) pGC2-sgRNA4 phage lysates (selection with with Kan (150 µg/ml) and Cm (10 µg/ml)), resulting in strains: (i) NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA1; (ii) NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA2; (iii) NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA3; (iv) NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA4, each expressing dCas9-sfGFP protein and the respective sgRNA. COL MurJ-mCherry pCNX-dCas9-GFP was transduced with previously prepared pGC2-MurJ sgRNA phage lysate (selection with Kan (150 µg/ml) and Cm (10 µg/ml)) to obtain COL MurJ-mCherry pCNX-dCas9-GFP pGC2-MurJ sgRNA.

Each plasmid (pGC2-sgRNA1 to sgRNA4) was also transduced into NCTC8325-4 FP650-RodZ (selection with Cm (10 µg/ml)), generating the control strains: (i) NCTC8325-4 FP650-RodZ pGC2-sgRNA1; (ii) NCTC8325-4 FP650-RodZ pGC2-sgRNA2; (iii) NCTC8325-4 FP650-RodZ pGC2-sgRNA3; (iv) NCTC8325-4 FP650-RodZ pGC2-sgRNA4. MurJ sgRNA was also transduced into COL MurJ-mCherry (selection with Cm (10 µg/ml)), resulting in control strain COL MurJ-mCherry pGC2-MurJ sgRNA.

Additional control strains were obtained by transduction of both pGC2 and pCNX empty vectors into COL MurJ-mCherry and NCTC8325-4 FP650-RodZ to obtain COL MurJ-mCherry pGC2 pCNX and NCTC8325-4 FP650-RodZ pGC2 pCNX. pGC2 was also transduced to COL MurJ-mCherry pCNX-dCas9-GFP, originating strain COL MurJ-

mCherry pCNX-dCas9-GFP pGC2. The transductants were selected with Kan (150 µg/ml) and Cm (10 µg/ml).

The constructed CRISPR strains were observed by fluorescence microscopy as described above. For that, NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA (sgRNA1 to sgRNA4), COL MurJ-mCherry pCNX-dCas9-GFP pGC2-MurJ sgRNA and the respective negative controls were grown in TSB supplemented with Kan (5 µg/ml), Cm (1 µg/ml) and Cd (0.1 µM for COL strains and 0.2 µM for the remaining cultures) until OD=0.4, before microscopy observation.

**Growth analysis of *S. aureus* strains**

Overnight cultures of COL MurJ-mCherry pCNX-dCas9-GFP pGC2-MurJ sgRNA, COL MurJ-mCherry pCNX-dCas9-GFP pGC2 and COLpMUTINsacol1804i were diluted to $OD_{600nm}$ 0.02 into fresh TSB with Kan (150 µg/ml) and Cm (10 µg/ml). Cd (0.1 µM) or IPTG (0.05 mM) were added when necessary and incubated at 37ºC with aeration. Growth was analysed by measuring, at regular intervals, the $OD_{600nm}$ of the cultures.

**Table 2. Bacterial strains used in this study**

| Strain | Relevant characteristics (a) | Source/Reference |
|---|---|---|
| ***E. coli*** | | |
| DC10B | Δdcm in the DH10B background; Dam methylation only | 17 |
| ***S. aureus*** | | |
| RN4220 | MSSA strain, restriction negative | 87 |
| NCTC8325-4 | MSSA strain | R. Novick |
| NCTC8325-4 FP650-RodZ | NCTC8325-4 expressing RodZ fused to fluorescent protein FP650 | H. Veiga and M. G. Pinho, unpublished |
| COL | MRSA strain | 88 |
| COL MurJ-mCherry | COL expressing MurJ fused to fluorescent protein mCherry | V. Macário and M. G. Pinho, unpublished |

| Strain | Relevant characteristics (a) | Source/Reference |
|---|---|---|
| COLpMUTINsacol1804i | COL expressing inducible MurJ under Pspac promoter | T. Ferreira and M. G. Pinho, unpublished |
| NCTC8325-4 FP650-RodZ pGC2-sgRNA1 | NCTC8325-4 expressing both RodZ fused to fluorescent protein FP650 and sgRNA1, $Cm^R$ | This study |
| NCTC8325-4 FP650-RodZ pGC2-sgRNA2 | NCTC8325-4 expressing both RodZ fused to fluorescent protein FP650 and sgRNA2, $Cm^R$ | This study |
| NCTC8325-4 FP650-RodZ pGC2-sgRNA3 | NCTC8325-4 expressing both RodZ fused to fluorescent protein FP650 and sgRNA3, $Cm^R$ | This study |
| NCTC8325-4 FP650-RodZ pGC2-sgRNA4 | NCTC8325-4 expressing both RodZ fused to fluorescent protein FP650 and sgRNA4, $Cm^R$ | This study |
| NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP | NCTC8325-4 expressing RodZ fused to fluorescent protein FP650 and *S. aureus* dCas9 fused to sfGFP, $Kan/Neo^R$ | This study |
| NCTC8325-4 FP650-RodZ pGC2 pCNX | NCTC8325-4 expressing RodZ fused to fluorescent protein FP650 and carrying pGC2 pCNX, $Cm^R$; $Kan/Neo^R$ | This study |
| COL MurJ-mCherry pGC2-MurJ sgRNA | COL expressing both MurJ fused to fluorescent protein mCherry and its sgRNA; $Cm^R$ | This study |
| COL MurJ-mCherry pCNX-dCas9-GFP | COL expressing both MurJ fused to fluorescent protein mCherry and *S. aureus* dCas9 fused to sfGFP; $Kan/Neo^R$ | This study |
| COL MurJ-mCherry pGC2 pCNX | COL expressing both MurJ fused to fluorescent protein mCherry and carrying pGC2 pCNX, $Cm^R$; $Kan/Neo^R$ | This study |
| COL MurJ-mCherry pCNX-dCas9-GFP pGC2 | COL expressing both MurJ fused to fluorescent protein mCherry and *S. aureus* dCas9 fused to sfGFP and carrying pGC2; $Cm^R$; $Kan/Neo^R$ | This study |
| NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA1 | NCTC8325-4 reporter strain expressing both dCas9 fused to sfGFP and sgRNA1; $Cm^R$, $Kan/Neo^R$ | This study |
| NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA2 | NCTC8325-4 reporter strain expressing both dCas9 fused to sfGFP and sgRNA2; $Cm^R$, $Kan/Neo^R$ | This study |
| NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA3 | NCTC8325-4 reporter strain expressing both dCas9 fused to sfGFP and sgRNA3; $Cm^R$, $Kan/Neo^R$ | This study |

| Strain | Relevant characteristics (a) | Source/Reference |
|---|---|---|
| NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP pGC2-sgRNA4 | NCTC8325-4 reporter strain expressing both dCas9 fused to sfGFP and sgRNA4; $Cm^R$, $Kan/Neo^R$ | This study |
| COL MurJ-mCherry pCNX-dCas9-GFP pGC2-MurJ sgRNA | COL reporter strain expressing both dCas9 fused to sfGFP and MurJ sgRNA; $Cm^R$, $Kan^R$ | This study |

(a) Pcad promoter – cadmium inducible Pcad promoter; Pspac promoter – IPTG inducible promoter; $Cm^R$ – chloramphenicol resistance; $Kan/Neo^R$ – kanamycin and neomycin resistant; sfGFP – super fast green fluorescent protein

**Table 3. Plasmids used in this study**

| Plasmid | Relevant Characteristics(a) | Source/Reference |
|---|---|---|
| pKILL FloA | KI vector expressing FloA-mNeongreen and *cat* between up- and downstream regions of *floA;* $Cm^R$, $Amp^R$ | M. Bramkamp, unpublished |
| pKILL TatA | KO vector expressing *cat* between up- and downstream regions of *tatA;* $Cm^R$, $Amp^R$ | M. Bramkamp, unpublished |
| pEPSA5 | *E. coli – S. aureus* shuttle vector; replicative in *S. aureus*; $Amp^R$, $Cm^R$; XylR | 82 |
| pTRC99a-P7 | Vector containing *p7superfastgfp*, $Amp^R$ | 83 |
| pCNX | Shuttle vector containing a Pcad promoter; $Amp^R$, $Kan/Neo^R$ | 84 |
| pGC2 | *E. coli/S. aureus* shuttle vector; $Amp^R$, $Cm^R$ | 85 |
| pCNX-GFP | pCNX encoding sfGFP under Pcad control; $Amp^R$, $Kan/Neo^R$ | This study |
| pCNX-dCas9-GFP | pCNX encoding *S. aureus* dCas9 fused to sfGFP under Pcad; $Amp^R$, $Kan/Neo^R$ | This study |
| pGC2-sgRNA1 | pGC2 encoding sgRNA1 (non ATG strand of FP650-RodZ) under *pbpb* promoter control; $Amp^R$, $Cm^R$ | This study |
| pGC2-sgRNA2 | pGC2 encoding sgRNA2 (ATG strand of FP650-RodZ) under *pbpb* promoter control; $Amp^R$, $Cm^R$ | This study |
| pGC2-sgRNA3 | pGC2 encoding sgRNA3 (ATG strand of FP650-RodZ) under J23119 promoter control; $Amp^R$, $Cm^R$ | This study |
| pGC2-sgRNA4 | pGC2 encoding sgRNA4 (ATG strand of FP650-RodZ) under J23119 promoter control; $Amp^R$, $Cm^R$ | This study |

| Plasmid | Relevant Characteristics(a) | Source/Reference |
|---|---|---|
| pGC2-MurJ sgRNA | pGC2 encoding MurJ sgRNA (ATG strand of MurJ-mCherry) under J23119 promoter control; $Amp^R$, $Cm^R$ | This study |

(a) Pcad promoter – cadmium inducible Pcad promoter; $Amp^R$ – ampicillin resistance; J23119 promoter – synthetic constitutive minimal promoter; $Cm^R$ – chloramphenicol resistance; $Kan/Neo^R$ – kanamycin and neomycin resistant; KI – knock-in; KO – knock-out; *cat* – chloramphenicol resistance gene; XylR - Xylose inducible pT5X promoter; sfGFP – super fast green fluorescent protein

**Table 4. Primers used in this study**

| Abbreviation | Primer | Sequence (5'-3')(b) | Source/Reference |
|---|---|---|---|
| P1 | NCTCTatA999Rv | GTTATGAACCATTGATAGCAC | This study |
| P2 | NCTCTatA999Fw | CCTAAAATTCCACCTATG | This study |
| P3 | NWMN FloA999 Rv | AACATCTTTGCATTGACTTAC | M. Bramkamp |
| P4 | NWMN FloA999 Fw | GTTTACGTGCAGCTTCT | M. Bramkamp |
| P5 | Amp_SeqP4Fw | GGTAGCTCTTGATCCGGC | This study |
| P6 | Amp_seq_P2 | CAGAAACGCTGGTGAAAG | H. Veiga |
| P7 | pKILLseqCAT_Rv | CATGCAGGATTGTTTATGAAC | This study |
| P8 | pKILLseqCAT_Fw | CTCTTTTCTCTTCCAATTGTC | This study |
| P9 | NucP1 | GCGATTGATGGTGATACGGTT | 89 |
| P10 | NucP2 | AGCCAAGCCTTGACGAACTAAAGC | 89 |
| P11 | GFPlinkerP1_Bam | CGTGGATCCTCCTGCGGCGCCTCCAGTAAAGGAGAAGAACTTTTCACTG | This study |
| P12 | GFPSTOPP2_Eco | CGCCGAATTCTTATTTGTATAGTTCATCCATGCC | This study |
| P13 | NCTCPBP2PromoterSmaI_P1Fw | CGCGCCCGGGCAACACATACTTGTACTTGCC | This study |
| P14 | sgRNAFP650EcoRI_P2Rv | GCGGGAATTCGTCATATGTAGAGAGGTACCTCGAGCGGCCCAAGCTTAAAAAAATCTCGCCAACAAGTTGACG | This study |

| Abbreviation | Primer | Sequence (5'-3')(b) | Source/Reference |
|---|---|---|---|
| | | AGATAAACACGGCATTTT GCCTTGTTTTAGTAGATTC TGTAATTTTCATTACAGAG TACTAAAAC**TCTAATTTA CGATCAACAAAA**TGCGAA AGACACAATACACC | |
| P15 | T7promoterprimer | TAATACGACTCACTATAG GG | P. Reed |
| P16 | gRNAFP650_P1F w | GTCTTTCGCA**AAAAGATT GTTTAAAGAAATC**GTTTT AGTACTCTGTAATGA | This study |
| P17 | gRNAFP650_P2Rv | GTACTAAAAC**GATTTCTT TAAACAATCTTTT**TGCGA AAGACACAATACACC | This study |
| P18 | sgRNAtermEcoRI_ Rv | GCGGGAATTCGTCATATGT AGAGAGGTACC | This study |
| P19 | SmaI_gRNA3FP65 0P1Fw | CGCGCCCGGGTTGACAGC TAGCTCAGTCCTAGGTATA ATACTAGT**AGAAACCAAA ACCACCACCAA** GTTTTAGTACTCTGTAATG A | This study |
| P20 | SmaI_gRNA4RodZ P1Fw | CGCGCCCGGGTTGACAGC TAGCTCAGTCCTAGGTATA ATACTAGT**GTCCCATTCG GCTTGGTTCGA**GTTTTAG TACTCTGTAATGA | This study |
| P21 | SmaI_gRNAMurJP 1Fw | CGCGCCCGGGTTGACAGC TAGCTCAGTCCTAGGTATA ATACTAGT**ATTTCTTTAC TTTCACTCATT**GTTTTAGT ACTCTGTAATGA | This study |
| P22 | murJm_P1 | CGGAGATCTTGGTATTACT GGAGAAAGCGTAAGC | 84 |
| P23 | murJm_P2 | ACCGACTGCACCGCTTCGT AAAAACCTAACTCTACG | 84 |
| P24 | murJm_P3 | AGCGGTGCAGTCGGTATG ATTGTGAGCAAGG | 84 |
| P25 | murJm_P4 | CGCCCATGGCTTGTACAGC TCGTCCATG | 84 |
| P26 | murJm_P5 | TGCCCATGGCGTAGAGTT AGGTTTTTACG | 84 |
| P27 | murJm_P6 | CGCGGATCCATACTTTCGT CAGTCTTATTG | 84 |

| Abbreviation | Primer | Sequence (5'-3')(b) | Source/Reference |
|---|---|---|---|
| P28 | FP650RodZ_P5_SalI | GCTGCGCTGTCGACGCTCATACATGCGCCTATTTATATC | H. Veiga |
| P29 | FP650RodZ_P6 | CTTCACCCATTTCATAGCCTCCTTACACTTACTCG | H. Veiga |
| P30 | FP650RodZ_P7 | GGAGGCTATGAAATGGGTGAAGATAGTGAATTAATTAG | H. Veiga |
| P31 | FP650RodZ_P2 | CAAGGAGGCGCCGCAGGAACTATGACCTAATTTTGATGG | H. Veiga |
| P32 | FP650RodZ_P3 | CATAGTTCCTGCGGCGCCTCCTTGAAAACGGTCGGTGAAGCGC | H. Veiga |
| P33 | FP650RodZ_P8_NcoI | TGCATTCCATGGCCATATGAAAATGAAATTAG | H. Veiga |

(a) Restriction sites are underlined in primers' sequence. gRNA sequences in bold.

# 3. Results and discussion

## pKILL plasmid as genetic tool for *S. aureus*

pKILL plasmid was designed by M. Bramkamp to be used as a vector to knock in (KI) or knock out (KO) DNA fragments in bacterial genomes through allelic replacement in a single step. Double crossover events between homologous regions in the genome and pKILL plasmid can be easily selected because single crossover events lead to the constitutive expression of toxin gene *yqdb,* which will kill bacterial cells.

Two plasmids were constructed in the Bramkamp lab, in the backbone of pKILL: pKILL TatA and pKILL FloA.

pKILL TatA was designed as a knock out vector. This plasmid carries homologous regions upstream and downstream of *tatA* gene in *S. aureus* Newman strain with the chloramphenicol resistance gene (*cat)* in between, to chromosomally replace *tatA* by *cat*. Hence, the KO mutant would be resistant to chloramphenicol and present a mutant phenotype resulting from the absence of *tatA*. *tatA* gene encodes for a translocase protein responsible by secreting many exo-enzymes, such as lecithinase, which degrades lecithin present in egg yolk. [90] Recombinants lacking *tatA* will not be able to degrade lecithin in egg yolk agar, therefore not showing a degradation halo in this medium.

On the contrary, pKILL FloA was designed as a KI vector. This plasmid carries a homologous fragment corresponding to the downstream flanking region of *floA* in the chromosome, and a fragment that comprises *cat*, with its corresponding promoter, the fluorescent marker *mNeongreen* and the 5' end of *floA* gene sequences. Therefore, double crossover with this vector results in the replacement of *floA* gene for *floA-mNeongreen* fusion. Consequently, successful KI colonies would be resistant to chloramphenicol and would express floA (a flotilin protein present in the cell membrane) fused to mNeongreen, observable by fluorescence microscopy as a single green fluorescent focus at the membrane.[91,92] Recombination processes for pKILL TatA and pKILL FloA are represented in Figure 15.

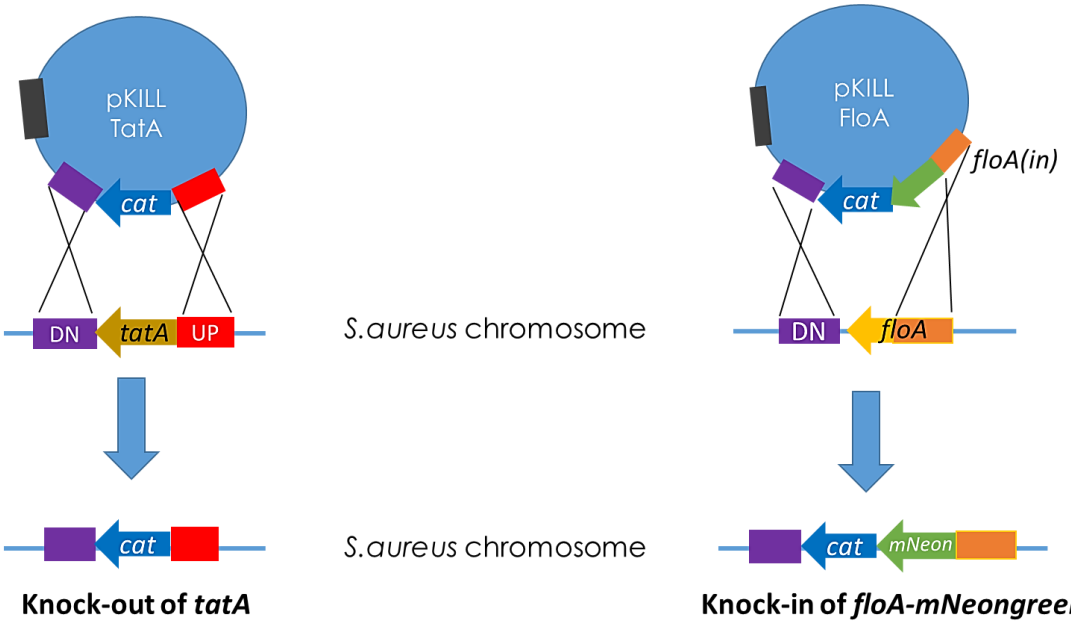



**Figure 15. Recombination events using pKILL TatA and pKILL FloA.** On the left, a double crossover event between pKILL TatA and *S. aureus* genome is shown. The homologous regions, upstream and downstream of *tatA*, are represented by red and violet boxes respectively. Recombination between those regions result in *S. aureus* mutants with genomic *tatA* gene replaced with *cat*, represented by a blue arrow. On the right, a double crossover between pKILL FloA and *S. aureus* genome. The violet box represents the downstream region of *floA* gene, whereas the orange box *floA(in)* represents the 5'end of *floA* and the green arrow, *mNeongreen.* Recombination between those regions leads to *S. aureus* mutants expressing *floA-mNeongreen* and *cat*. In both plasmids, the grey box represents the toxin encoding gene *yqdb.*

Having two available plasmids, pKILL FloA and pKILL TatA, our main goal was to test whether a pKILL vector is useful for *S. aureus* genome editing.

In order to define the best conditions to test pKILL plasmids, we first performed a round of test electroporations with pKILL TatA and pKILL FloA at different selective antibiotic concentrations. In this document we only present results corresponding to experiments following the final optimizations, which were performed according to the general workflow shown in Figure 16. Briefly, we introduced each of the pKILL plasmids into RN4220 competent cells and confirmed if the electroporated colonies were the expected mutants, presenting the KI or KO, by phenotypic and genotypic analysis - which was performed by PCR using primers complementary to the chromosomal upstream and downstream regions of *tatA* and *floA.* In all electroporations, a negative control (RN4220 electroporated without plasmid) and a positive control (RN4220 electroporated with 0.12

µg pEPSA5) were used, in order to confirm if the electroporations were correctly performed and if neither the plates nor the competent cells were contaminated.

For each electroporation experiment, and in order to increase the probability of having a high number of recombinant mutants, we used high concentrations of plasmid DNA (5.6-17 µg). Transformation efficiency of RN4220 cells was tested for each batch by electroporation with 1 µl (120 µg/ml) of a high-copy number plasmid, pEPSA5. The competent cells with higher transformation efficiency were used for electroporation.

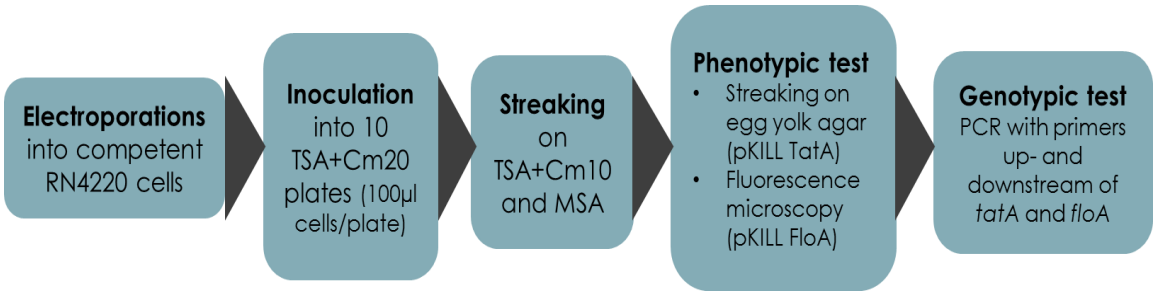



**Figure 16. General workflow for screening of *S. aureus* mutants generated using pKILL plasmid.**

### Use of pKILL TatA did not result in *tatA* knock-out *S. aureus* mutants

Following the workflow presented in Figure 16, three electroporations were made with pKILL TatA into RN4220 competent cells: experiment A, B and C with 6.8, 13 and 17 µg DNA electroporated respectively. Together, these experiments generated a total of 100 mutant colonies (Table 5), from which 98 were probably *S. aureus* (as considered from their mannitol degradation, shown by a yellow halo on MSA plates) and were resistant to Cm. The remaining two colonies did not produce the expected yellow halo on MSA plates, and so were not considered for the subsequent phenotypic/genotypic analysis.

The 98 mutants produced a degradation halo on egg yolk agar plates similar to the wild-type control (5 examples in Figure 17). 50 of these 98 mutants were tested by colony screening PCR with primers P1/P2 upstream and downstream of *tatA* gene, which showed that none of the colonies presented the expected *tatA* KO. Like the negative control, RN4220, all tested colonies presented a 2kb band that corresponds to the expected amplification without KO (11 of the 50 colonies tested are presented in Figure 18).

According to our results, it was not possible to obtain recombinant RN4220 pKILL TatA colonies. The presence of chloramphenicol resistant colonies in the absence of *tatA* KO leads to relevant questions regarding the origin of their resistance, being thus important to further analyze their genotypes in order to determine how their antibiotic susceptibility was altered.

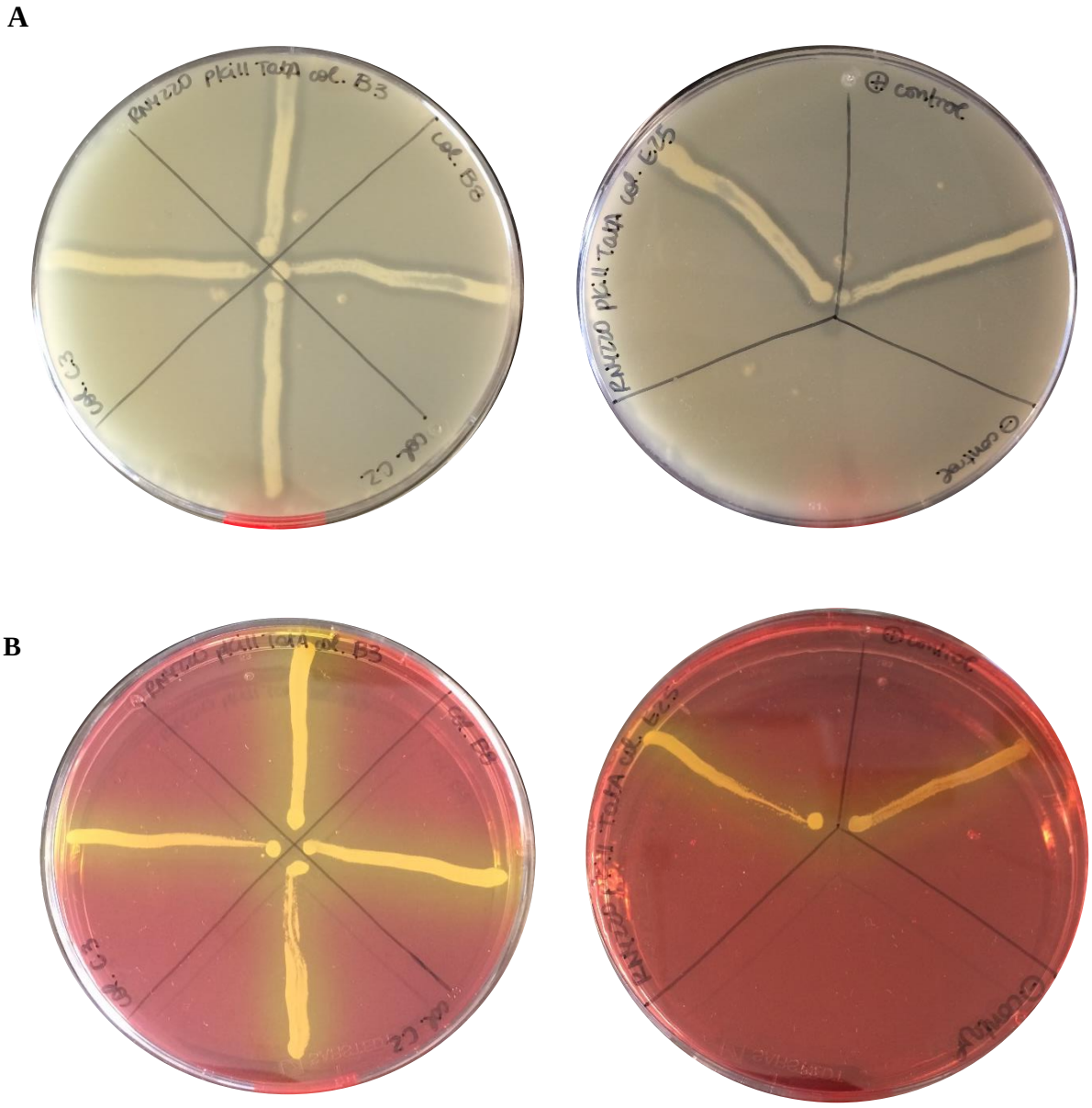



**Figure 17. Phenotypic analysis for 5 RN4220 pKILL TatA colonies on egg yolk agar and MSA plates.** A: RN4220 pKILL TatA colonies B3, B8, C2, C3 and E25 on egg yolk agar, with RN4220 pEPSA and RN4220 as plate controls. The degradation halo in every colony is visible. B: RN4220 pKILL TatA colonies B3, B8, C2, C3 and E25 on MSA plates with RN4220 and DC10B as plate controls. The colonies were considered to be probably *S. aureus*, due to the growth in MSA with the characteristic yellow halo.

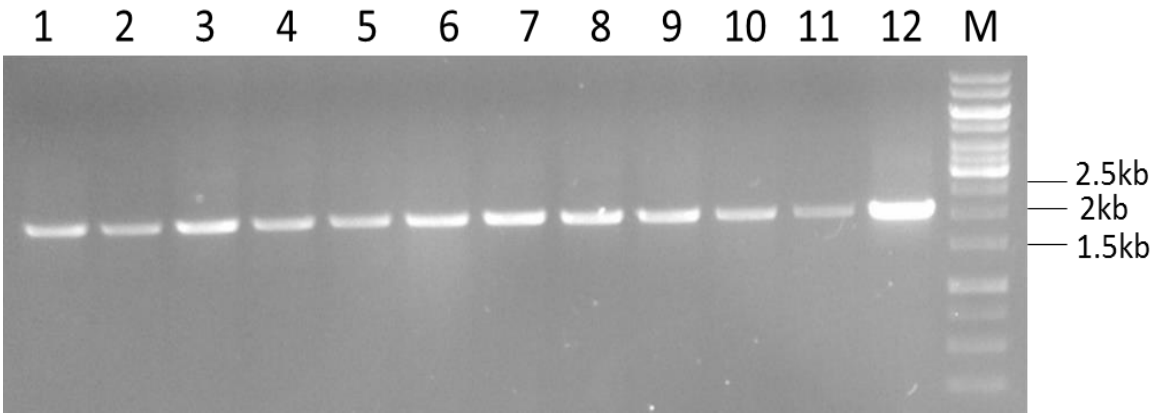


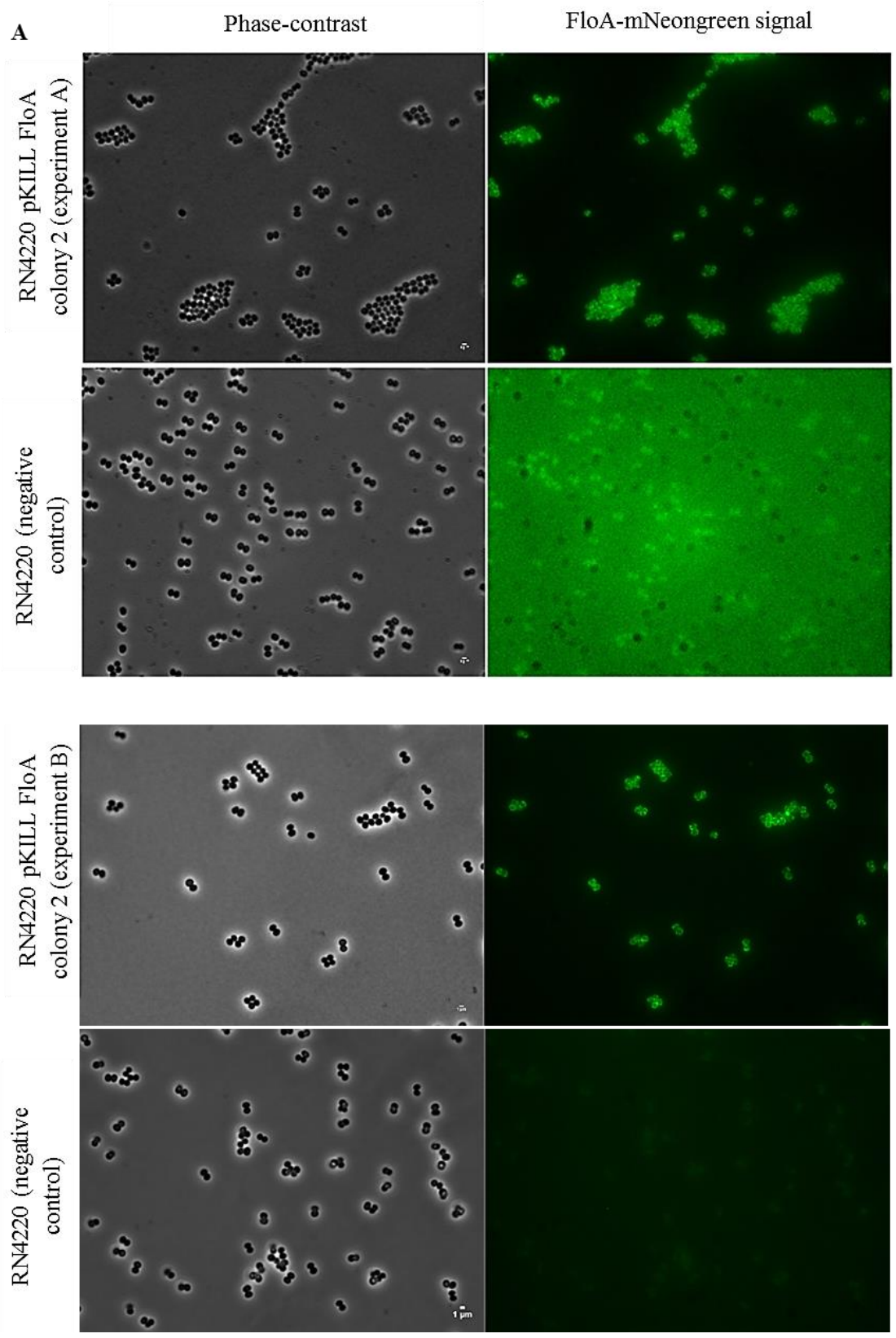

**Figure 18. DNA electrophoresis gel of the PCR products obtained from the screening of 11 RN4220 pKILL TatA electroporation colonies with primers P1/P2.** Lanes 1-11 correspond to 11 electroporation colonies, out of 50 total colonies; Lane 12 corresponds to the negative control RN4220 and Lane M corresponds to the molecular marker. All tested colonies present the same 2kb band as the negative control, which corresponds to the wild-type fragment without KO. The expected KO band would correspond to a 2.5kb fragment. Electrophoresis performed in 0.8% agarose with 4 µl PCR samples and 3 µl 1kb DNA Ladder.

**Table 5. Summary of pKILL TatA electroporations**

| | Experiment | | |
|---|---|---|---|
| | A | B | C |
| Competence (transformation efficiency) | 30 transformants/ng DNA | | 70 transformants/ng DNA |
| Electroporated DNA (µg) | 6.8 | 13 | 17 |
| Number of electroporation colonies obtained | 18 | 42 | 40 |
| Yellow degradation halo on MSA | 18 | 42 | 38 |
| Resistant to Cm | 18 | 42 | 38 |
| Halo in egg yolk agar | 18 | 42 | 38 |
| KO genetic confirmation (in 50 tested colonies) | 0 | 0 | 0 |

**pKILL FloA was successfully used to generate *S. aureus* strains expressing *floA-mNeongreen***

To test pKILL FloA, two different electroporations were made: experiment A and B, in which 5.6 and 6.0 μg of DNA were electroporated, respectively (Table 6). A total of 9 RN4220 mutants were obtained, from which 8 produced a yellow degradation halo on MSA plates, as expected for *S. aureus*, and were resistant to Cm (Table 6). Through fluorescence microscopy observation of these 8 colonies, it was possible to verify that only 7 had *floA-mNeongreen* signal at the membrane, whereas the remaining colony had no fluorescence at the membrane, as seen on Figure 19A and B respectively. These results showed that allelic replacement with pKILL FloA had occurred, resulting in chloramphenicol resistant RN4220 cells expressing FloA-mNeongreen.


A
Phase-contrast
FloA-mNeongreen signal
RN4220 pKILL FloA colony 2 (experiment A)
RN4220 (negative control)
RN4220 pKILL FloA colony 2 (experiment B)
RN4220 (negative control)
1 µm

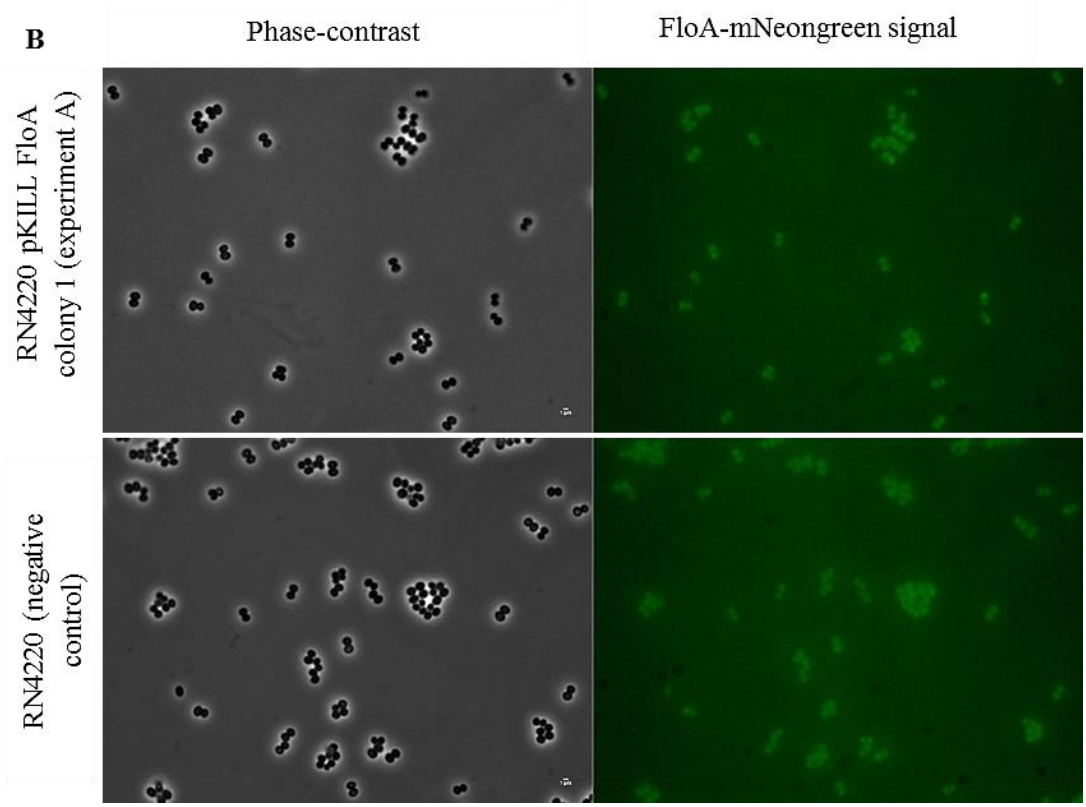



**Figure 19. RN4220 cells expressing *floA-mNeongreen* constructed using pKILL FloA**. Epifluorescence microscopy images of RN4220 transformants that resulted from pKILL FloA electroporations. A: Two different colonies, producing FloA-mNeongreen signal, from two electroporations are presented. Underneath the fluorescent colonies are the corresponding negative controls, RN4220, observed on the same day. B: A colony from experiment B does not show FloA-mNeongreen signal. Underneath is the corresponding negative control, RN4220. Panels show phase-contrast images (left) and FloA-mNeongreen fluorescence, false-colored (right).

All the colonies that presented FloA-mNeongreen signal were submitted to colony screening PCR with primers P3/P4 which hybridize upstream and downstream of *floA* locus (Table 4), to confirm the KI. 3 colonies among the tested fluorescent mutants presented a band that corresponds to the KI (~4.3kb); whereas the remaining 4, from experiment B, did not have the expected amplified KI band (Figure 20).

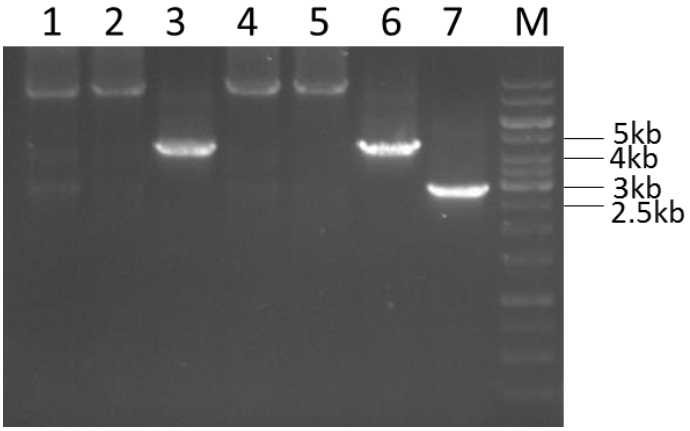



**Figure 20. DNA electrophoresis gel of the PCR products obtained from the screening of pKILL FloA RN4220 electroporation B colonies with primers P3/P4.** Lanes 1-6 correspond to electroporation colonies 1-5, 7 respectively; Lane 7 corresponds to the negative control RN4220. Lane M corresponds to the molecular marker. The negative control shows the expected 3kb band whereas colonies 3 and 7 have a ~4.3kb band corresponding to the KI. The remaining colonies present bands that do not correspond to the expected genotype. Electrophoresis performed in 0.8% agarose with 4 µl PCR samples and 3 µl 1kb DNA Ladder.

To try to understand why colonies that present antibiotic resistance and *floA-mNeongreen* expression did not have the expected *floA-mNeongreen* KI, we re-tested them with different sets of primers. If only one homologous recombination event had occurred, the cells could have pKILL FloA integrated into the genome, carrying in this way the ampicillin resistance gene, along with *mNeongreen*, *cat,* and *yqdb.* Therefore, we used primers P5/P6 (Table 4) complementary to the ampicillin resistance gene (*amp*), to screen the colonies by PCR for the presence of *amp* gene. All colonies tested presented a 1kb fragment confirming the presence of *amp*, as seen on Figure 21A. However, with this PCR we could not conclude if *amp* was present because the plasmid was integrated into the genome or was still present at the cytoplasm. To test which one of these hypothesis was correct, we used the pairs of primers P7/P3 and P8/P4 (see Table 4), which include a primer complementary to Newman genome (P3 and P4), and a primer complementary to *cat* gene in the plasmid (P7 and P8); that allows us to detect if a single cross-over event had occurred (amplification of a ~7kb fragment, see Figure 22). This PCR, however, presented inconclusive results with unspecific bands of amplification (Fig. 21B); including bands that actually correspond to a KI when using primers P7/P3 and P8/P4 (1kb and 3kb bands respectively) rendering impossible to determine the origin of those mutants' resistance or to obtain the KI confirmation.

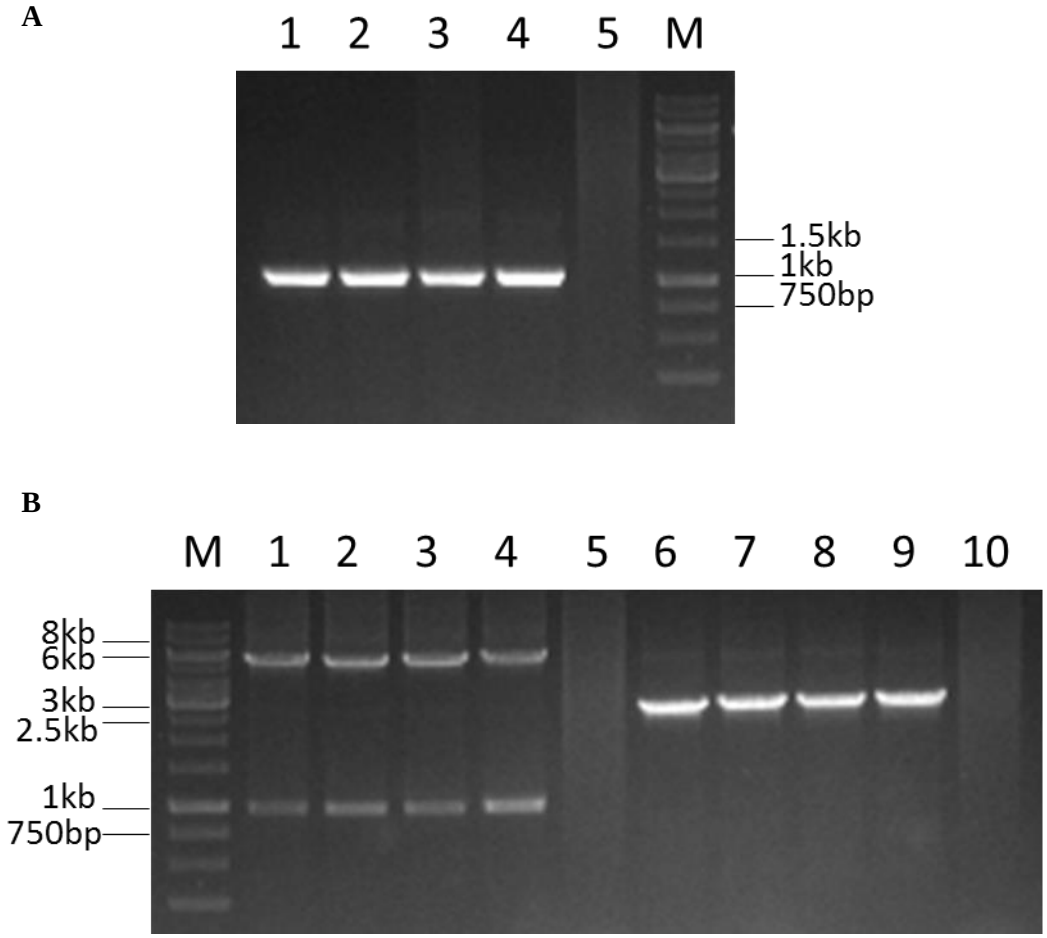



**Figure 21. DNA electrophoresis gels of RN4220 pKILL FloA without confirmed KI.** A: DNA electrophoresis gel of the PCR products obtained from pKILL FloA RN4220 electroporation colonies with primers P5/P6 (Lane 1-4). Lanes 1-4 correspond to electroporation colonies 1, 2, 4, 5 respectively; Lane 5 corresponds to the negative control RN4220. All colonies, except the negative control, present the 1.1kb band corresponding to amplification of a fragment in *amp* gene. B: DNA electrophoresis gel of the PCR products obtained from pKILL FloA RN4220 electroporation colonies with primers P3/P7 (Lane 1-4) and P4/P8 (Lane 6-9). Lanes 1-4, 6-9 correspond to electroporation colonies 1, 2, 4, 5 respectively; Lane 10 corresponds to the negative control RN4220. Lane 1-4 present both 1 and 5-6kb bands that do not correspond to the expected single cross-over band with primers P3/P7 (7kb). Lane 6-9 present a unique 3kb band, which does not equally correspond to the expected amplification with P4/P8. 1kb and 3kb bands with each pair of primers could correspond ideally to the KI band with those primers. In both gels, Lane M corresponds to the molecular marker. Electrophoresis performed in 0.8% agarose with 4 µl PCR samples and 3 µl 1kb DNA Ladder.

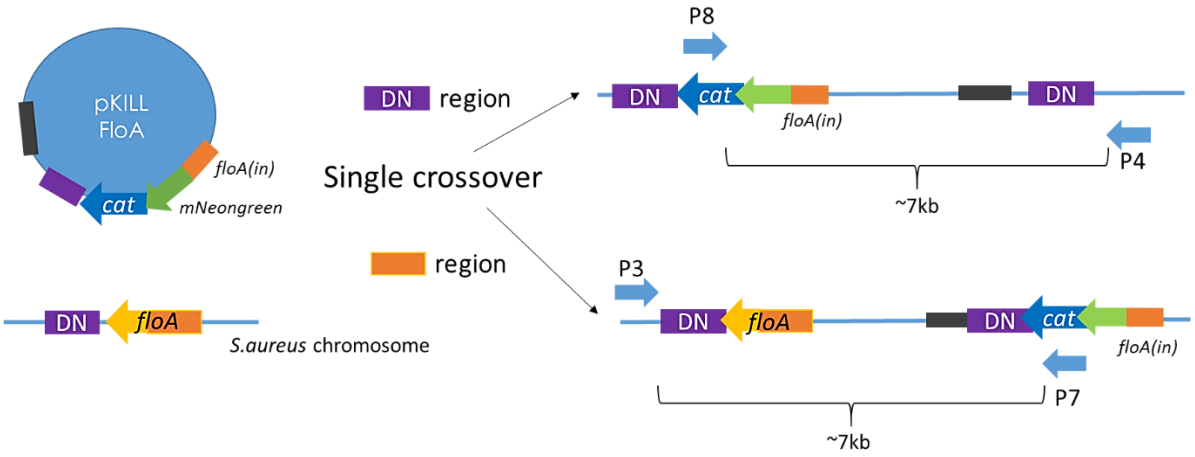



**Figure 22. Schematic representation of a putative pKILL FloA single cross-over.** pKILL FloA is represented as the blue circle with *yqdb (*grey rectangle) and homologous regions to the genome corresponding to the downstream (violet rectangle) and 5'end regions of *floA* (orange rectangle). The blue and green arrows between those homologous regions in the plasmid represent *cat* and *mNeongreen* respectively. In the eventuality of a single crossover event, the genome would rearrange through downstream or 5'end homologous regions as shown on the right. Primers used to confirm plasmid integration through single crossover are represented as the blue arrows underneath each genomic rearrangement, amplifying approximately 7kb fragments in each case.

It should be noticed that even if pKILL FloA is indeed incorporated into the genome, or if it is present in the cytoplasm, it is still difficult to explain why the cells seem to be resistant to the expression of the toxin codified by *yqdb*. Possibly by mutating this gene or its promoter, the cells could have found a mechanism to survive and propagate.

**Table 6. Summary of pKILL FloA electroporations**

| | **Experiment** | |
|---|---|---|
| | A | B |
| Competence (transformation efficiency) | $1.5x10^2$ transformants/ng DNA | $2x10^2$ transformants/ng DNA |
| Electroporated DNA (µg) | 5.6 | 6.0 |
| Number of electroporation colonies obtained | 2 | 7 |
| Yellow degradation halo on MSA | 2 | 6 |
| Resistant to Cm | 2 | 6 |
| FloA-mNeongreen signal | 1 | 6 |
| KI genetic confirmation | 1 | 2 |
| Single cross-over | 0 | Inconclusive |
| Resistance to ampicillin | 0 | 4 |

Despite the likely low efficiency of recombination (as shown by the low number of confirmed RN4220 pKILL FloA colonies), these results indicate that is possible to perform genome editing in highly competent RN4220 cells by using pKILL FloA, proving that using pKILL to achieve allelic replacement in one single step in *S. aureus* is possible. However, the fact that we did not obtain any confirmed *tatA* KO RN4220 mutants with pKILL TatA shows that gene editing with pKILL system is not achieved at a consistent level. Moreover, although the obtained colonies for pKILL FloA and pKILL TatA were resistant to antibiotic and expressed *floA-mNeongreen* (for RN4220 pKILL FloA), not all had the desired KI/KO, leading to unanswered questions regarding the origin of such phenotypes that would have to be investigated before future applications. Therefore, we concluded that although allelic replacement through pKILL plasmid is functional in *S. aureus* cells, its low efficiency and lack of reproducibility proves the need for further optimization before establishing pKILL vectors as a standard tool for the study of *S. aureus*.

## CRISPRi as a genome editing tool for *S. aureus*

Our goal was to construct a CRISPRi system for *S. aureus* cells that could be useful for the deletion and study of essential genes. For that, we started by developing and optimizing the components of this system, using a non-essential gene reporter. We chose to use a NCTC8325-4 strain expressing a RodZ N-terminal fusion to a far-red fluorescent protein, eqFP650 (H. Veiga and M. G. Pinho, unpublished). RodZ, a protein responsible for cell shape in *Caulobacter crescentus*[93], was previously shown in our Laboratory to be non-essential in *S. aureus* and to localize to the septum (Figure 23). A NCTC8325-4 Δ*rodZ* mutant does not have obvious phenotypical defects under the microscope and presents the same growth rate as the wild-type strain. In NCTC8325-4 FP650-RodZ strain, the genomic *rodZ* was replaced by *fp650-rodZ*, making this fusion the only RodZ copy expressed in the cell. This strain can be used to easily test the efficacy of a CRISPR system designed to control gene expression, as *fp650-rodZ* depletion can be directly observed by fluorescence microscopy.

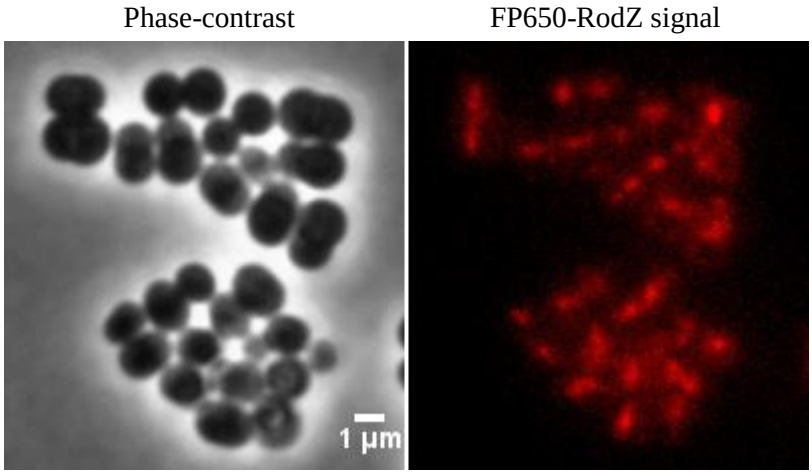



**Figure 23. FP650-RodZ localizes at the septum.** Epifluorescence microscopy images of *S. aureus* NCTC8325-4 FP650-RodZ strain that expresses FP650-RodZ fusion as the only RodZ copy in the cells. Panels show, from left to right, phase-contrast images and FP650-RodZ fluorescence, false-colored.

In this study, we chose to use Sa dCas9 instead of the widely used Cas9 from *Streptococcus pyogenes* (SpCas9) since Sa Cas9 is a native protein of *S. aureus* and so, most likely to be more efficient than other Cas9s and less prone to interfere with *S. aureus* cell metabolism. According to our knowledge, this Cas9 was never applied to genome editing of prokaryotes, which increases its testing interest. In order to be able to quickly detect SaCas9 expression, we fused it to GFP. Since eqFP650 has a different emission

peak than GFP, it's possible to observe the expression of these two proteins simultaneously by fluorescence microscopy.[94]

The last component required for the CRISPRi system is the sgRNA, the guiding component required for specific Cas9 binding, and thus, the determining factor of CRISPR editing success.[42] In this study, we determined the best sgRNA targeting region for efficient gene repression and selected the ideal promoter to control its expression.

**sgRNA designed to target the ATG strand is highly efficient in repressing gene expression**

To optimize the sgRNA sequence, and determine which strand allowed for more efficient repression, we designed two different sgRNAs that target the two strands of *fp650-rodZ* DNA sequence in NCTC8325-4 FP650-RodZ: sgRNA1 and sgRNA2.

The first step in designing the sgRNAs was to select the correct gRNA sequence. For proper target recognition by the Cas9 – sgRNA complex, the gRNA should be complementary to a sequence located at the 5' end of a protospacer adjacent motif (PAM) in the non-target strand.[39] Each Cas9 recognizes a specific PAM. SaCas9 recognizes 5'-NNGRRT-3' PAM.[62] For sgRNA1 design, we selected a 21-nt sequence (gRNA1) that is complementary to the 5'end region of the non ATG strand of *fp650* (Target1), located next to a Sa Cas9 PAM on the ATG strand as shown in Figure 24. The non ATG strand is defined as the non-coding strand, and is read by the RNA polymerase. sgRNA2 was designed by selecting a 21-nt gRNA2 sequence that is complementary to the middle of the ATG strand of *fp650* (Target2), immediately next to a PAM on the non ATG strand (Figure 24). For both sgRNAs, we used 21nt as it was shown to be an ideal gRNA length.[37]

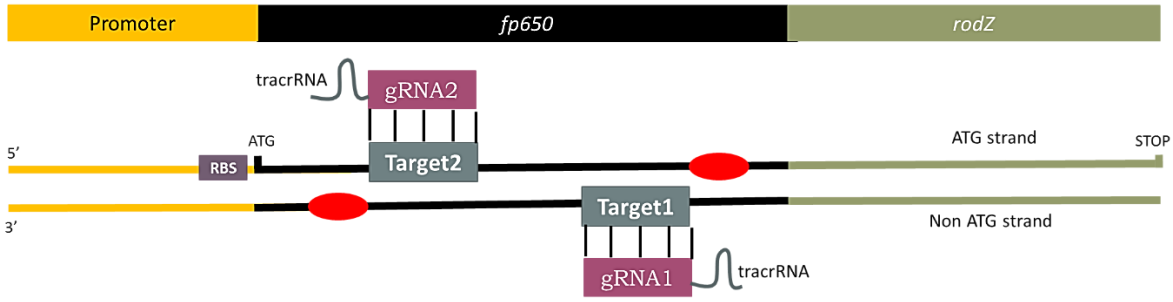



**Figure 24. Schematic representation of sgRNAs to control *fp650-rodZ* expression.** The two DNA strands of *fp650-rodZ* gene are represented, including its promoter region. sgRNA1 is composed of a 21nt gRNA complementary to the 5' end region of *fp650* in the non ATG strand, fused to a tracrRNA. sgRNA2 is complementary to the middle region of *fp650* in the ATG strand. The required PAM sequences for each sgRNA are represented as a red circle and are found next to the respective target sequence in the non-target strand.

Besides selecting the gRNA sequence, the choice of promoter was also an important component for successful sgRNA expression. To express constitutively sgRNA1 and sgRNA2 in *S. aureus* cells, we chose the previously described *S. aureus* native *pbpb* promoter.[86] The *pbpb* promoter region consists in fact of three promoters that express either PBP2 alone or PBP2 with RecU, depending on which promoter is activated. The first of such promoters, P1[86], was chosen to express the sgRNA, since it presents higher activity than the other promoters and has a defined +1 (transcription start). A defined +1 site is important for the sgRNA design since the expression of more nucleotides in addition to the gRNA sequence was reported to decrease Cas9 target binding.[52] Both sgRNA1 and sgRNA2 were expressed from the high-copy number vector pGC2. pGC2-sgRNA1 and pGC2-sgRNA2 vectors were introduced into NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP cells and the final strains were observed by fluorescence microscopy.

As shown in Figure 25A, after the introduction of CRISPRi system with sgRNA1, it is still possible to detect the expression of FP650-RodZ at the septum, much like the control strain without CRISPRi system - showing that sgRNA1 is inefficient for *fp650-rodZ* repression. The CRISPRi system with sgRNA2, on the other hand, showed a high efficiency of repression in comparison with sgRNA1, leading to lack of FP650-RodZ signal at the septum, simultaneously with dCas9-GFP expression (Figure 25B).

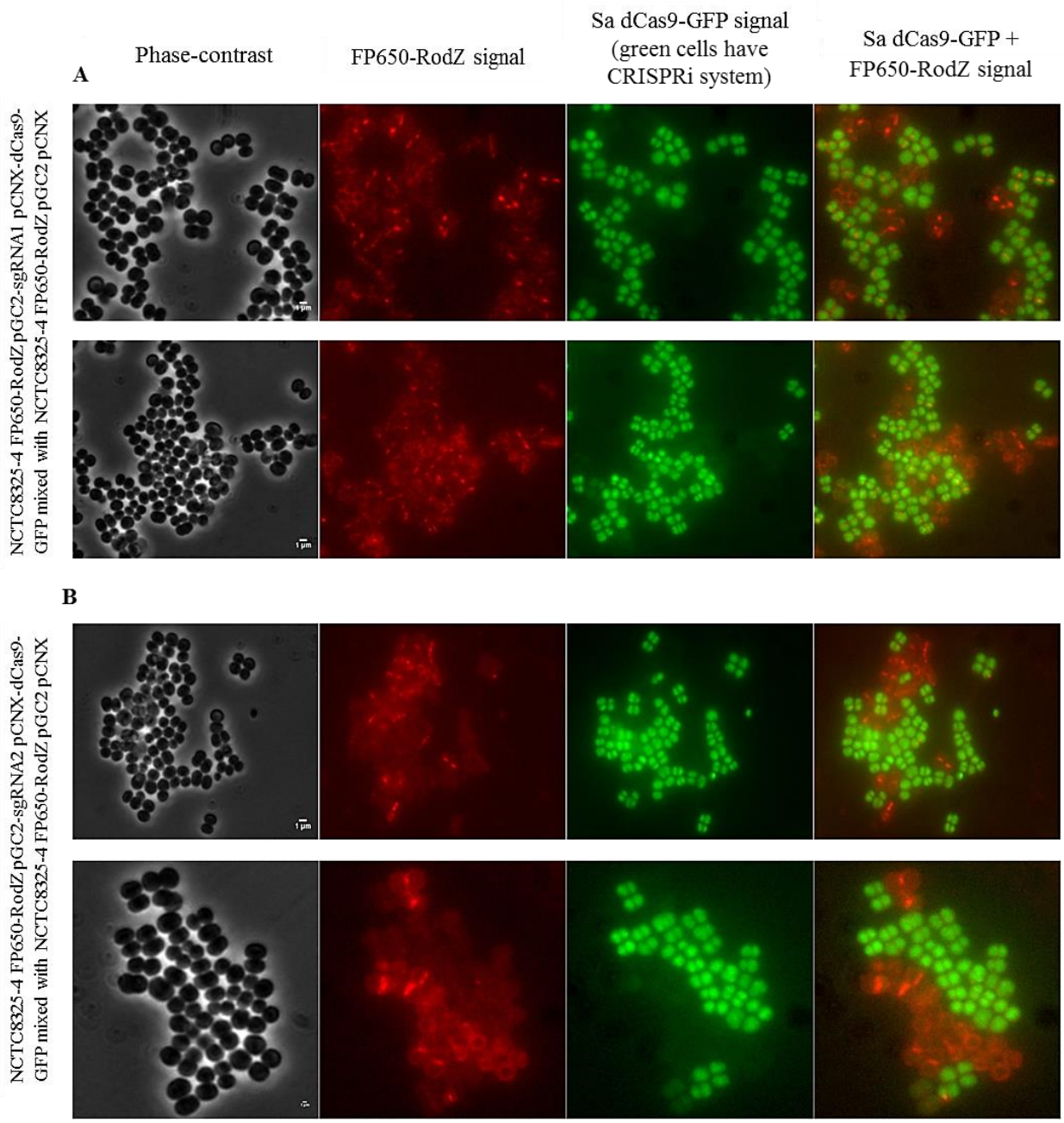



**Figure 25. Targeting ATG strand is more efficient than targeting non ATG strand.** A: NCTC8325-4 FP650-RodZ pGC2-sgRNA1 pCNX-dCas9-GFP was mixed with NCTC8325-4 FP650-RodZ pGC2 pCNX (negative control) and observed in the same slide. Both cultures are distinguished by the presence of GFP signal, which is observed only in NCTC8325-4 FP650-RodZ pGC2-sgRNA1 pCNX-dCas9-GFP strain. Notice that green cells (NCTC8325-4 FP650-RodZ pGC2-sgRNA1 pCNX-dCas9-GFP) still present FP650-RodZ signal (red), similarly to the wild-type, or negative, control (NCTC8325-4 FP650-RodZ pGC2 pCNX). B: NCTC8325-4 FP650-RodZ pGC2-sgRNA2 pCNX-dCas9-GFP was mixed with the negative control (NCTC8325-4 FP650-RodZ pGC2 pCNX) and observed in the same slide. Both cultures are distinguished by the presence of GFP signal, which is observed only in NCTC8325-4 FP650-RodZ pGC2-sgRNA2 pCNX-dCas9-GFP strain. Notice that green cells (NCTC8325-4 FP650-RodZ pGC2-sgRNA2 pCNX-dCas9-GFP) do not have FP650-RodZ (red) signal. All images were false-colored. Both strains were induced with 0.02µM cadmium before microscopy, to induce dCas9-GFP expression.

In order to confirm this effect was not due to any other component other than the CRISPRi system, and that both dCas9 and sgRNA2 individually were not able to alter FP650-RodZ signal, we observed by fluorescence microscopy the strains NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP and NCTC8325-4 FP650-RodZ pGC2-sgRNA2, that express only dCas9 or sgRNA2, respectively. As shown in Figure 26, the expression of dCas9 or sgRNA2 individually did not repress *fp650-rodZ* expression.

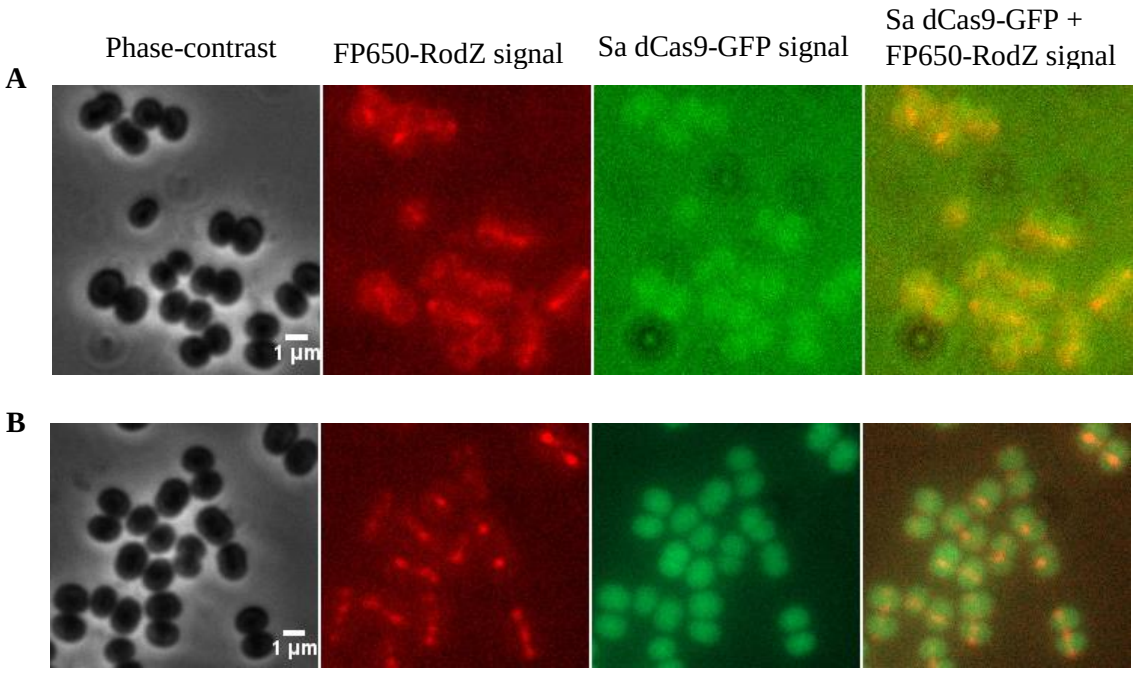



**Figure 26. dCas9 and sgRNA2 individually do not repress FP650-RodZ expression.** A: NCTC8325-4 FP650-RodZ pGC2-sgRNA2. B: NCTC8325-4 FP650-RodZ pCNX-dCas9-GFP, grown with 0.02 µM cadmium to induce dCas9-GFP expression. Both strains present FP650-RodZ signal at the septum. All images were false-colored.

In conclusion, these results show that the ATG strand is the best target for an efficient gene repression. This is in accordance with previous studies, which also tested both strands for sgRNA design and have reported the ATG strand as the most efficient target, although it is not the strand to which RNA polymerase binds.[68,95] In parallel to the results we have obtained, it was also proven using NCTC8325-4 FP650-RodZ cells that the best targeting strand for Sp dCas9 is the ATG strand (M. Sorg and M. G. Pinho, unpublished).

Further confirmation of CRISPRi repression of FP650-RodZ was not achieved by growth analysis since *rodZ* is a non-essential gene, therefore its repression would not

cause major alterations in cell growth that could be used for CRISPRi assessment. However, other options besides growth rate assays might include studying FP650-RodZ protein levels by Western-Blot or mRNA levels by qRT-PCR (quantitative real-time PCR), providing a quantitative and more detailed analysis of the actual result of CRISRi application in *S. aureus* cells.

### CRISPRi induces repression independently of the targeting region

After proving that the CRISPRi system that we designed works, and to further optimize it, we tested if the rate of repression efficiency was altered by targeting different regions of the gene of interest.

To this end, we designed an experiment, in the same reporter strain, using sgRNA2 and new sgRNAs: sgRNA3 and sgRNA4, which are complementary to the promoter and the 3'end regions of *fp650-rodZ*, respectively (Figure 27). Both sgRNAs have the same structure as sgRNA2, carrying a 21-nt gRNA fused to a scaffold, and are expressed by the constitutive minimal promoter J23119, which has a defined +1 start and a sequence much shorter than *pbpb* promoter's. This synthetic promoter is not so susceptible as *pbpb* promoter to be potentially controlled by *S. aureus* transcriptional regulators and has been successfully used for CRISPR purposes in other studies as well as in our Laboratory (data not shown).[68]


**Figure 27. sgRNA3 and sgRNA4 binding to *fp650-rodZ*.** Representation of *fp650-rodZ* gene sequence including its promoter region and corresponding RBS (purple box). The sgRNA3 is composed of a 21nt gRNA complementary to the ATG strand promoter region of *fp650-rodZ*, fused to a tracrRNA. The sgRNA4 binds to the 3'end region of *fp650-rodZ*. Each target PAM sequence is highlighted by a red circle. sgRNA3 and sgRNA4 were also used to test the efficiency of transcriptional inhibition using Sp dCas9 (data from M. Sorg and M. G. Pinho, unpublished). The PAM sequences for Sp dCas9 are highlighted by a blue circle.

sgRNA3 and sgRNA4 were cloned into a pGC2 vector, under the control of J23119 promoter. Each sgRNA vector was introduced into NCTC8325-4 FP650-RodZ

pCNX-dCas9-GFP cells to assess subsequent *fp650-rodZ* repression. The resulting strains were observed by fluorescence microscopy, using NCTC8325-4 FP650-RodZ pGC2 pCNX as negative control. In all the microscopy observations that were performed, from which two selected crops are shown (Figure 28A and B), it was possible to observe that the CRISPRi cells expressing sgRNA3 or sgRNA4 respectively did not show FP650-RodZ signal, while expressing dCas9-GFP. These results indicate that both the target to the promoter or the 3'end of the gene are effective in repressing *fp650-rodZ*, and at similar levels than sgRNA2, which targets the middle of *fp650* in *fp650-rodZ* DNA sequence.

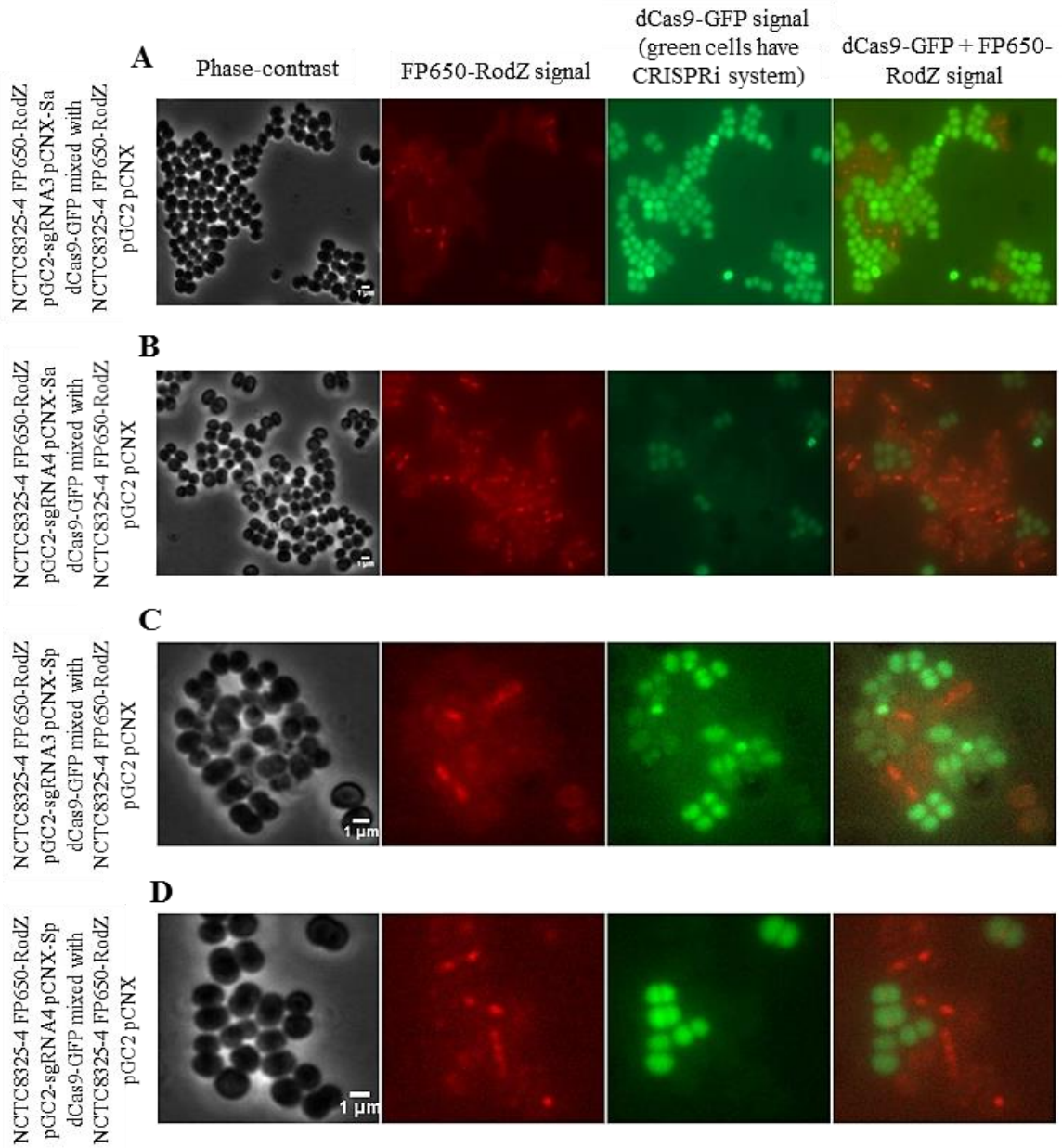



**Figure 28. sgRNA3 and sgRNA4, along with dCas9, induce repression of *fp650-rodZ*.** A: NCTC8325-4 FP650-RodZ pGC2-sgRNA3 pCNX-dCas9-GFP was mixed and observed with NCTC8325-4 FP650-RodZ pGC2 pCNX on the same slide. Both cultures are distinguished by the presence of GFP signal, seen

in NCTC8325-4 FP650-RodZ pGC2-sgRNA3 pCNX-dCas9-GFP strain. B: NCTC8325-4 FP650-RodZ pGC2-sgRNA4 pCNX-dCas9-GFP was mixed and observed with NCTC8325-4 FP650-RodZ pGC2 pCNX on the same slide. Both cultures are distinguished by the presence of GFP signal, seen in NCTC8325-4 FP650-RodZ pGC2-sgRNA4 pCNX-dCas9-GFP strain. A and B express Sa dCas9. Notice in A and B that green cells (expressing the CRISPRi system) do not present FP650-RodZ signal (red). C: NCTC8325-4 FP650-RodZ pGC2-sgRNA3 pCNX-Sp dCas9-GFP was mixed and observed with NCTC8325-4 FP650-RodZ pGC2 pCNX on the same slide. Both cultures are distinguished by the presence of GFP signal, seen in NCTC8325-4 FP650-RodZ pGC2-sgRNA3 pCNX-Sp dCas9-GFP strain. D: NCTC8325-4 FP650-RodZ pGC2-sgRNA4 pCNX-Sp dCas9-GFP was mixed and observed with NCTC8325-4 FP650-RodZ pGC2 pCNX on the same slide. Both cultures are distinguished by the presence of GFP signal, seen in NCTC8325-4 FP650-RodZ pGC2-sgRNA4 pCNX-Sp dCas9-GFP strain. C and D express Sp dCas9. Notice in C and D that green cells (expressing the CRISPRi system with Sp dCas9) do not present FP650-RodZ signal (red). All strains were grown in 0.02 µM cadmium to induce dCas9-GFP expression. From left to right (for A to D): phase-contrast images; cells submitted to eqFP650 excitation filter; cells submitted to sfGFP excitation filter; overlay of both fluorescent signals. All images were false-colored.

Taking into account the results obtained by fluorescence microscopy, we concluded that CRISPRi repression is effective independently of the targeting region of the ATG strand, presenting comparable rates of transcription inhibition when targeting the promoter, the middle of *fp650* or 3'end regions of *fp650-rodZ.* These results differ from previous studies in *E. coli* that reported the promoter as the most effective target for dCas9 repression.[56,68] It would therefore be interesting to proceed with quantitative analysis of gene repression using the CRISPRi system that we designed, in order to confirm that its efficacy is independent of the targeting region. We also concluded that the minimal J23119 constitutive promoter can be used for sgRNA expression, improving on the *pbpb* promoter as it is shorter and not controlled by native *S. aureus* transcriptional regulators. Finally, we compared these results with Sp dCas9, which was simultaneously being tested in our laboratory by M. Sorg with specific sgRNAs, expressed from J23119 promoter and targeting the same regions as sgRNA3 and sgRNA4. Both sgRNA3 and sgRNA4 were designed by us to target regions proximal to Sa and Sp Cas9 PAMs, allowing us to directly compare the efficiency of Sa and Sp dCas9s (Figure 27). Therefore, NCTC8325-4 FP650-RodZ strains expressing Sa dCas9 and sgRNA3 or sgRNA4 were observed by fluorescence microscopy in parallel with NCTC8325-4 FP650-RodZ strains expressing Sp dCas9 and sgRNAs that target the same regions (see Figure 28C and D). As for Sa dCas9, the results obtained with Sp dCas9 indicate that both targeting regions induce efficient repression. (M. Sorg and M. G. Pinho, unpublished) Moreover, we also observed that both nucleases presented repression at similar levels, as was previously reported by Ran et al. for DNA cleavage.[53] (M. Sorg and M. G. Pinho, unpublished)

Therefore, we concluded that both Sa and Sp dCas9 are effective in gene repression, independently of the targeting region.

### Application of CRISPRi to the repression of essential genes

All the results presented so far showed that the CRISPRi system we have developed is a useful tool to repress non-essential *fp650-rodZ* gene. It is our goal to test if it can be used for repression of essential genes.

For that, we designed a sgRNA to target the 5'end of the essential gene MurJ (SACOL1804), a putative lipid II flippase that localizes at the septum and is required for peptidoglycan synthesis.[84] Using a reporter strain that expresses MurJ-mCherry as the only MurJ copy in the cell, we transduced Sa dCas9-GFP and the designed sgRNA under J23119 promoter control, into its cells to assess CRISPRi repression. 206 colonies were grown in the transduction plates without Cd, to lower dCas9 expression since dCas9 is under the control of Cd inducible Pcad promoter. However, when analyzing the growth of the resulting strain in medium supplemented with Cd, we observed that culture growth rate was not altered (Figure 29). This growth curve suggested that control of dCas9 was not tight since the strain growth rate with and without inducer was similar. In this same experiment, the positive control was strain COLpMUTINsacol1804i, in which MurJ expression is induced by the presence of IPTG and whose growth rate decreased significantly without inducer. Moreover, it was possible to observe by fluorescence microscopy (Figure 30) that cells with the CRISPRi system presented a visible dCas9-GFP signal. This was even visible in the absence of cadmium (data not shown), showing that our system was leaky for dCas9 expression.

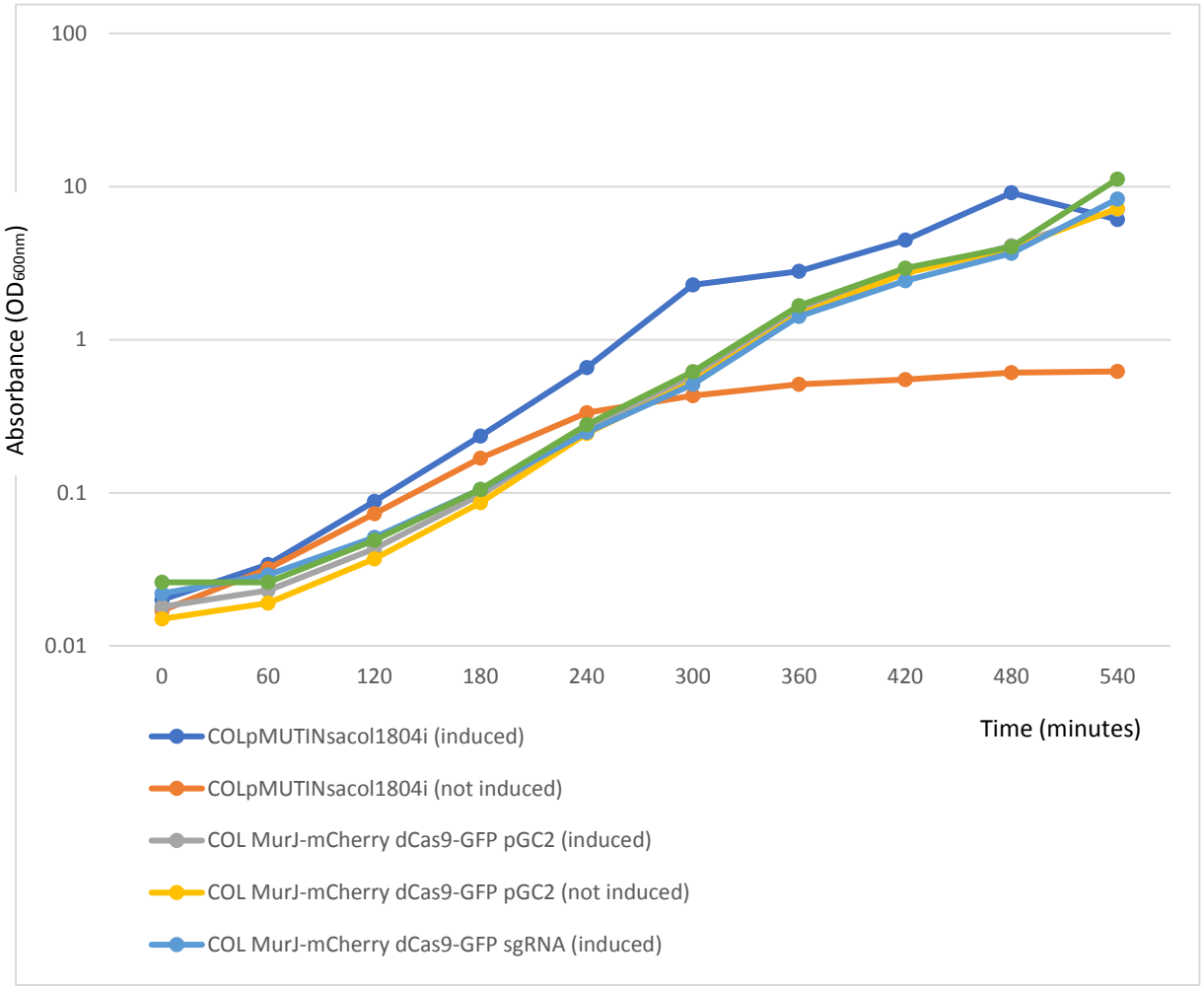



**Figure 29. *S. aureus* cells expressing dCas9-GFP and a MurJ sgRNA have the same growth rate as a wild-type strain.** Growth curves of COL MurJ-mCherry pCNX-dCas9-GFP pGC2-MurJ sgRNA strain, induced and not induced with Cd and the respective controls: COLpMUTINsacol1804i as the positive control, induced and not induced with IPTG; and COL MurJ-mCherry pCNX-dCas9-GFP pGC2, induced and not induced with Cd.

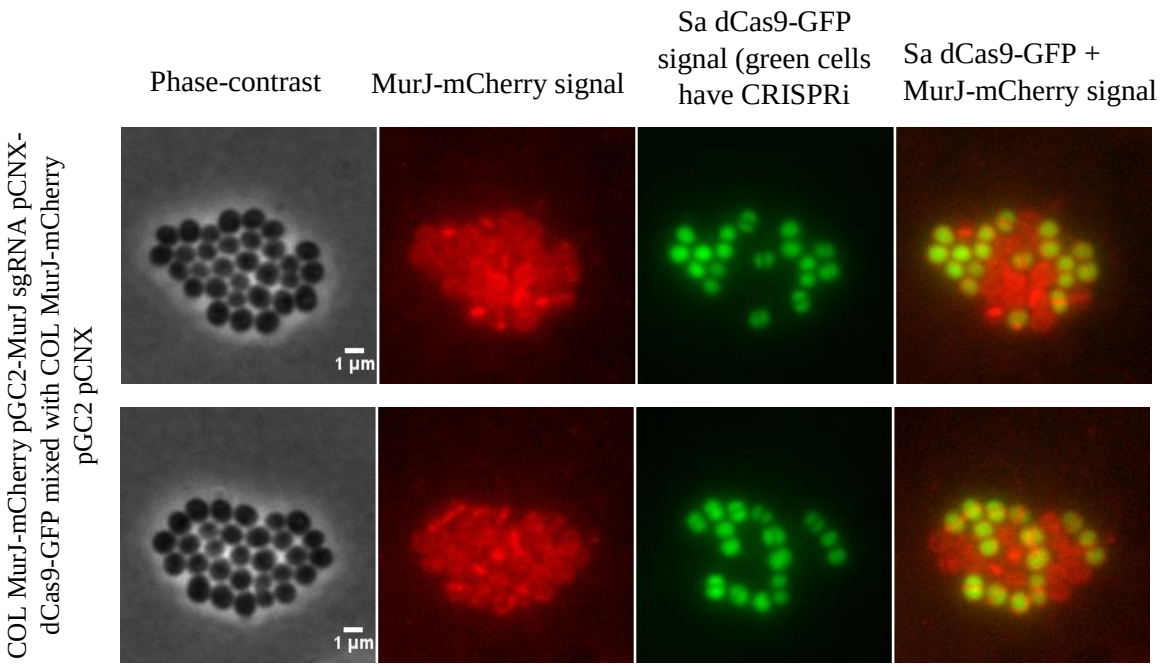



**Figure 30. *S. aureus* cells expressing dCas9-GFP and MurJ sgRNA maintain *murJ-mCherry* expression.** Two crops of images taken with COL MurJ-mCherry pCNX-dCas9-GFP pGC2-MurJ sgRNA, which was mixed and observed with COL MurJ-mCherry pCNX pGC2 in the same slide. Both strains are distinguished by the presence of GFP signal, found in COL MurJ-mCherry pCNX-dCas9-GFP pGC2-MurJ sgRNA strain. Notice that green cells (expressing the CRISPRi system) present MurJ-mCherry signal (red). All strains were grown in 0.01 µM cadmium to induce dCas9-GFP expression. From left to right: phase-contrast images; MurJ-mCherry fluorescence signal; dCas9-GFP signal; overlay of both fluorescent signals. All images were false-coloured.

These unexpected results suggest that the system we have designed, though inducible, might not be sufficiently controlled to be useful to repress essential genes. In fact, the Pcad promoter (present in pCNX) was previously implied to be leaky, and thus, this system's inability to repress *murJ* might be explained by the leaky expression of dCas9 that, by causing lethal conditions due to permanent repression of an essential gene, would result in selection of suppressor mutations (in dCas9 sequence, the sgRNA or their corresponding promoters). Therefore, for future CRISPRi applications (especially to repress essential genes), it is very important to design a system that is effectively induced and not leaky. For that, we suggest inducible systems using different promoters than Pcad that are more tightly controlled.

# 4. Conclusion

Our ability to genetically manipulate *S. aureus* is limited by the small number of tools available for this purpose. Here, we tested two new promising techniques designed to allow a fast, easy and reliable manipulation of *S. aureus* genome: the allelic replacement plasmid pKILL and CRISPRi.

According to the results for the different pKILL plasmids tested, we verified that this system allows allelic replacements as desired, but at a very low efficiency level, requiring extremely competent RN4220 cells. The presence of colonies containing the pKILL antibiotic resistance marker but without KO/KI also leads us to consider the possibility of existing mutations, possibly in the toxin encoding gene, which would interfere with pKILL's efficacy and allow the cells to survive and propagate in such conditions. Therefore, we conclude that this system, although functional in *S. aureus* cells, is not efficient enough to consistently produce recombinant mutants.

We successfully developed a CRISPRi repression system based on the native Cas9 of *S. aureus*. In this system, sgRNAs targeting the sequence of interest are constitutively expressed from a high-copy number vector while an inactive Sa dCas9, fused to GFP, is expressed under the control of Pcad promoter.[63] The expression of dCas9-GFP, instead of dCas9 alone, allows us to confirm by fluorescence microscopy that the protein is effectively being expressed. The binding of the dCas9 to a target sequence, guided by the sgRNA, stops the progression or interaction of the RNA polymerase, impeding gene expression.[68]

For the construction of the CRISPRi system, it was necessary to obtain an optimized sgRNA. To achieve this purpose, we implemented a series of CRISPR experiments targeting the non-essential reporter gene, *fp650-rodZ*. We concluded that the sgRNA should target the ATG (coding) strand in any region of a gene of interest, since targeting the promoter, middle or 3'end regions presented similar rates of repression. Moreover, by comparing results obtained after choosing identical targets for Sa and Sp dCas9, we concluded that both dCas9s successfully repressed *fp650-rodZ* independently of the targeting region, presenting comparable levels of efficiency as previously described on literature.[53]

The CRISPRi system that we implemented is still not optimized to be a truly inducible system, since dCas9 expression is probably not sufficiently controlled for the repression of essential genes. During the experiments we made with CRISPRi targeting an essential gene, the presence of grown colonies without major alterations implied that this system, in particular dCas9 expression, might be too leaky, promoting selection of suppressor mutations and resistant cells' propagation. The solution is in the future to test other alternatives for dCas9 expression control.

# 5. Future Work

Although the tools we tested in this study are promising for *S. aureus* genetic manipulation, they also present limitations that should be accounted for. For both techniques, improvements are still required in order to achieve a fully efficient *S. aureus* genome manipulation.

In pKILL experiments, since we could not exactly determine the presence of pKILL plasmid (replicative or integrated by single crossover) in colonies resistant to the pKILL antibiotic marker but without KO/KI, it would be relevant to analyse how these cells have grown in such conditions. The resulting answers could help to elucidate how to improve pKILL plasmid for higher efficiency. Therefore, we suggest genotypical analysis of those colonies, such as whole genome sequencing, to detect suppressor mutations or the presence of integrated pKILL plasmid; and further experiments to determine the presence of the replicative plasmid, such as PCR with their plasmid DNA. We also suggest assessing *S. aureus* cells susceptibility to the Yqdb protein, in order to confirm whether this toxin is a proper selection marker for double crossover mutants.

The CRISPRi system that we successfully constructed and applied to the repression of non-essential genes, needs also to be further optimized in order to become a tightly inducible system for dCas9 that can be applied for repression of essential genes. For that, it is necessary to decrease or eliminate leaky dCas9-GFP expression by placing it under the control of a tighter inducible promoter, such as tetO promoter, regulated by TetR, and previously used in *S. aureus*.[95]

We also suggest introducing dCas9 into the chromosome, under an inducible promoter and expressing the gene from a single copy, to avoid expression heterogeneity (as verified by plasmid expression). As previously reported in gene screenings using CRISPR in *Bacillus subtilis* and *Streptococcus pneumoniae*, chromosomal dCas9 controlled by xylose-inducible Pxyl or IPTG-inducible lacI promoters respectively have been successfully applied to essential genes.[96,97] Chromosomal dCas9 expression, controlled by inducible promoters, does not completely avoid leakiness but decreases basal expression, proving to be an interesting alternative for inducible systems.

The development of a tightly controlled CRISPRi system would originate a promising tool for the study of *S. aureus*. Ideally, after attaining a fully optimized CRISPRi system, essential and non-essential genes could be targeted and repressed with similar efficiency, allowing for further insight into their roles. This application could be particularly interesting for important protein complexes in *S. aureus* such as the divisome, which is responsible for cell growth and division, and whose organization and assembly still requires deeper understanding. Although a comprehensive knowledge on *S. aureus* cell division mechanisms would be most welcomed, especially when antibiotic resistance is so widely distributed, manipulation of the divisome genes can be a difficult task since many of these proteins are encoded in complex operons and the available tools that can be used for that purpose are not always efficient. Therefore, practical and easy techniques such as CRISPRi that could be able to overcome these limitations are necessary for the genetic manipulation of *S. aureus*.

# 6. References


1. Harris Llinos G, Foster SJ, Richards RG, Lambert P, Stickler D, Eley A. An introduction to *Staphylococcus aureus*, and techniques for identifyingand quantifying *S. aureus* adhesins in relation to adhesion to biomaterials. *Eur Cells Mater*. 2002;4:39-60. doi:10.22203/eCM.v004a04.

2. Thomer Lena, Schneewind Olaf, Missiakas Dominique. Pathogenesis of *Staphylococcus aureus* Bloodstream Infections. *Annu Rev Pathol Mech Dis*. 2016;11(1):343-364. doi:10.1146/annurev-pathol-012615-044351.

3. Orenstein Abigail. The Discovery and Naming of *Staphylococcus aureus*. *Infect Dis Antimicrob agents*. 1998:1880-1881. doi:10.1021/bc049951i.

4. Tong Steven YC, Davis Joshua S, Eichenberger Emily, Holland Thomas L, Fowler Vance G. *Staphylococcus aureus* infections: Epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev*. 2015;28(3):603-661. doi:10.1128/CMR.00134-14.

5. Kazakova Sophia V, Hageman Jeffrey C, Matava Matthew, Srinivasan Arjun, Phelan Larry, Garfinkel Bernard, Boo Thomas, McAllister Sigrid, Anderson Jim, Jensen Bette, Dodson Doug, Lonsway David, McDougal Linda K, Arduino Matthew, Fraser Victoria J, Killgore George, Tenover Fred C, Cody Sara, Jernigan Daniel B. A Clone of Methicillin-Resistant *Staphylococcus aureus* among Professional Football Players. *N Engl J Med*. 2005;352(5):468-475. doi:10.1056/NEJMoa042859.

6. Muto CA, Jernigan JA, Ostrowsky BE, Richet HM, Jarvis WR, Boyce JM, Farr BM. SHEA guideline for preventing nosocomial transmission of multidrug- resistant strains of *Staphylococcus aureus* and *Enterococcus*. *Infect Control Hosp Epidemiol*. 2003;24(5):362-386. doi:10.1086/502213.

7. Miller LG, Diep BA. Colonization, Fomites, and Virulence: Rethinking the Pathogenesis of Community-Associated Methicillin-Resistant *Staphylococcus aureus* Infection. *Clin Infect Dis*. 2008;46(5):752-760. doi:10.1086/526773.

8. Boucher Helen W, Corey G Ralph, Boucher Helen W, Corey G Ralph. Epidemiology of Methicillin-Resistant *Staphylococcus aureus*. *Clin Infect Dis*. 2008;46(Supplement_5):S344-S349. http://dx.doi.org/10.1086/533590.

9. Barber Mary, Rozwadowska-Dowzenko Mary. Infection By Penicillin-Resistant Staphylococci. *Lancet*. 1948;252(6530):641-644. doi:10.1016/S0140-6736(48)92166-7.

10. Peacock Sharon J, Paterson Gavin K. Mechanisms of Methicillin Resistance in *Staphylococcus aureus*. *Annu Rev Biochem*. 2015;84(1):577-601. doi:10.1146/annurev-biochem-060614-034516.

11. Seligman Stephen J. Penicillinase-Negative Variants of Methicillin-Resistant *Staphylococcus aureus*. *Nature*. 1966;209(5027):994-996. http://dx.doi.org/10.1038/209994a0.

12. Jevons M Patrici, Coe AW, Parker MT. Methicillin resistance in Staphylococci. *Lancet*. 1963;281(7287):904-907. doi:https://doi.org/10.1016/S0140-6736(63)91687-8.

13. Lyon BR, Skurray R. Antimicrobial resistance of *Staphylococcus aureus*: genetic basis. *Microbiol Rev*. 1987;51(1):88-134.

14. Vogel Valérie, Falquet Laurent, Calderon-Copete Sandra P, Basset Patrick, Blanc Dominique S. Short term evolution of a highly transmissible methicillin-resistant *Staphylococcus aureus* clone (ST228) in a Tertiary care hospital. *PLoS One*. 2012;7(6):1-9. doi:10.1371/journal.pone.0038969.

15. LD Saravolatz, N Markowitz, L Arking, D Pohlod, E Fisher. Methicillin-resistant *Staphylococcus aureus*: Epidemiologic observations during a community-acquired outbreak. *Ann Intern Med*. 1982;96(1):11-16. doi:10.7326/0003-4819-96-1-11.

16. Monk Ian R. Genetic manipulation of Staphylococci—breaking through the barrier. *Front Cell Infect Microbiol*. 2012;2(April):1-9. doi:10.3389/fcimb.2012.00049.

17. Monk Ian R, Shah Ishita M, Xu Min. Transforming the Untransformable : Application of Direct. *MBio*. 2012;3(2):e00277-11. doi:10.1128/mBio.00277-11.Editor.

18. Prax Marcel, Lee Chia Y, Bertram Ralph. An update on the molecular genetics toolbox for Staphylococci. *Microbiol (United Kingdom)*. 2013;159(PART3):421-435. doi:10.1099/mic.0.061705-0.

19. Yansura Daniel G, Hennert Dennis J. Use of the *Escherichia coli* lac repressor and operator to control gene expression in *Bacillus subtilis* (hybrid promoter/isopropyl 13-D-thiogalactoside induction). *Biochemistry*. 1984;81(January):439-443. doi:10.1073/pnas.81.2.439.

20. Brantl Sabine. Antisense RNAs in plasmids: Control of replication and maintenance. *Plasmid*. 2002;48(3):165-173. doi:10.1016/S0147-619X(02)00108-7.

21. Nakashima Nobutaka, Miyazaki Kentaro. Bacterial cellular engineering by genome editing and gene silencing. *Int J Mol Sci*. 2014;15(2):2773-2793. doi:10.3390/ijms15022773.

22. Forsyth R Allyn, Haselbeck Robert J, Ohlsen Kari L, Yamamoto Robert T, Xu Howard, Trawick John D, Wall Daniel, Wang Liangsu, Brown-Driver Vickie, Froelich Jamie M, C. Kedar G, King Paula, McCarthy Melissa, Malone Cheryl, Misiner Brian, Robbins David, Tan Zehui, Zhu Zhan-yang, Carr Grant, Mosca Deborah A, Zamudio Carlos, Foulkes J Gordon, Zyskind Judith W. A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*. *Mol Microbiol*. 2002;43(6):1387-1400. doi:10.1046/j.1365-2958.2002.02832.x.

23. Ji Yinduo, Zhang Barbara, Van Stephanie F, Horn, Warren Patrick, Woodnutt Gary, Burnham Martin KR, Rosenberg Martin. Identification of Critical Staphylococcal Genes Using Conditional Phenotypes Generated by Antisense RNA. *Science (80- )*. 2001;293(5538):2266 LP-2269. doi:10.1126/science.1063566.

24. Vagner V, Dervyn E, Ehrlich SD. A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology*. 1998;144(11):3097-3104. doi:10.1099/00221287-144-11-3097.

25. Arnaud Maryvonne, Chastanet Arnaud, Débarbouillé Michel. New Vector for Efficient Allelic Replacement in Naturally Nontransformable, Low-GC-Content, Gram-Positive Bacteria. *Appl Env Microbiol*. 2004;70(11):6887-6891. doi:10.1128/AEM.70.11.6887-6891.2004.

26. Harwood Colin R, Pohl Susanne, Smith Wendy, Wipat Anil. Chapter 4 - *Bacillus subtilis*: Model Gram-Positive Synthetic Biology Chassis. In: Harwood Colin, Wipat Anil BT Methods in Microbiology, eds. *Microbial Synthetic Biology*. Vol 40. Academic Press; 2013:87-117. doi:https://doi.org/10.1016/B978-0-12-417029-2.00004-2.

27. Bae Dongryeoul, Seo Keun Seok, Zhang Ting, Wang Chinling. Characterization of a Potential *Listeria monocytogenes* Virulence Factor Associated with Attachment to Fresh Produce. *Appl Environ Microbiol*. 2013;79(22):6855-6861. doi:10.1128/AEM.01006-13.

28. Silvaggi Jessica M, Perkins John B, Losick Richard. Small untranslated RNA antitoxin in *Bacillus subtilis*. *J Bacteriol*. 2005;187(19):6641-6650. doi:10.1128/JB.187.19.6641-6650.2005.

29. Sorek Rotem, Kunin Victor, Hugenholtz Philip. CRISPR a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Micro*. 2008;6(3):181-186. doi:10.1038/nrmicro1793.

30. de la Fuente-Núñez César, Lu Timothy K. CRISPR-Cas9 technology: applications in genome engineering, development of sequence-specific antimicrobials, and future prospects. *Integr Biol*. 2017;9(2):109-122. doi:10.1039/C6IB00140H.

31. Marraffini Luciano A, Sontheimer Erik J. CRISPR interference limits horizontal gene transfer in Staphylococci by targeting DNA. *Science (80- )*. 2008;322(5909):1843-1845. doi:10.1126/science.1165771.

32. Marraffini Luciano A, Sontheimer Erik J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet*. 2010;11(3):181-190. doi:10.1038/nrg2749.

33. Horvath P, Barrangou R. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science (80- )*. 2010;327(5962):167-170. doi:10.1126/science.1179555.

34. Modell Joshua W, Jiang Wenyan, Marraffini Luciano A. CRISPR–Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature*. 2017;544(7648):101-104. doi:10.1038/nature21719.

35. Jinek Martin, Chylinski Krzysztof, Fonfara Ines, Hauer Michael, Doudna Jennifer A, Charpentier Emmanuelle. A Programmable Dual-RNA Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80- )*. 2012;337(6096):816-821. doi:10.1126/science.1225829.

36. CRISPR-Cas adaptive immune system. http://noobnim.in.th/crispr-cas-system/crispr-cas-adaptive-immune-system/. Accessed December 12, 2017.

37. Chen Baohui, Hu Jeffrey, Almeida Ricardo, Liu Harrison, Balakrishnan Sanjeev, Covill-Cooke Christian, Lim Wendell A, Huang Bo. Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res*. 2016;44(8):e75--e75. doi:10.1093/nar/gkv1533.

38. Briner Alexandra E, Barrangou Rodolphe. Guide RNAs: A glimpse at the sequences that drive CRISPR–Cas systems. *Cold Spring Harb Protoc*. 2016;2016(7):594-600. doi:10.1101/pdb.top090902.

39. Mohanraju Prarthana, Makarova Kira S, Zetsche Bernd, Zhang Feng, Koonin Eugene V, van der Oost John. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science (80- )*. 2016;353(6299). doi:10.1126/science.aad5147.

40. Choi Kyeong Rok, Lee Sang Yup. CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnol Adv*. 2016;34(7):1180-1209. doi:10.1016/j.biotechadv.2016.08.002.

41. CRISPR/Cas9 Transfection. https://www.mirusbio.com/applications/genome-editing-using-crispr-cas. Accessed December 12, 2017.

42. Jinek Martin, East Alexandra, Cheng Aaron, Lin Steven, Ma Enbo, Doudna Jennifer. RNA-programmed genome editing in human cells. *Elife*. 2013;2(2):e00471. doi:10.7554/eLife.00471.

43. Xu Han, Xiao Tengfei, Chen Chen-Hao Hao, Li Wei, Meyer Clifford A, Wu Qiu, Wu Di, Cong Le, Zhang Feng, Liu Jun S, Brown Myles, Liu X Shirley. Sequence determinants of improved CRISPR sgRNA design. *Genome Res*. 2015;25(8):1147-1157. doi:10.1101/gr.191452.115.

44. Cong Le, Ran F Ann, Cox David, Lin Shuailiang, Barretto Robert, Hsu Patrick D, Wu Xuebing, Jiang Wenyan, Marraffini Luciano A. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science (80- )*. 2013;339(6121):819-823. doi:10.1126/science.1231143.Multiplex.

45. Mali Prashant, Yang Luhan, Esvelt Kevin M, Aach John, Guell Marc, DiCarlo James E, Norville Julie E, Church George M. RNA-Guided Human Genome Engineering via Cas9. *Science (80- )*. 2013;339(6121):823-826. doi:10.1126/science.1232033.RNA-Guided.

46. Pattanayak Vikram, Lin Steven, Guilinger John P, Ma Enbo, Doudna Jennifer A, Liu David R. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotech*. 2013;31(9):839-843. http://dx.doi.org/10.1038/nbt.2673.

47. Synthetic single guide RNA for CRISPR-Cas9 experiments. http://dharmacon.gelifesciences.com/applications/gene-editing/crispr-cas9-genome-editing-synthetic-99mer-sgrna/. Accessed December 12, 2017.

48. Mohr Stephanie E, Hu Yanhui, Ewen-Campen Benjamin, Housden Benjamin E, Viswanatha Raghuvir, Perrimon Norbert. CRISPR guide RNA design for research applications. *FEBS J*. 2016;283:3232-3238. doi:10.1111/febs.13777.

49. Dang Ying, Jia Gengxiang, Choi Jennie, Ma Hongming, Anaya Edgar, Ye Chunting, Shankar Premlata, Wu Haoquan. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol*. 2015;16(1):280. doi:10.1186/s13059-015-0846-3.

50. Haeussler Maximilian, Concordet Jean Paul. Genome Editing with CRISPR-Cas9: Can It Get Any Better? *J Genet Genomics*. 2016;43(5):239-250. doi:10.1016/j.jgg.2016.04.008.

51. Zhang Jian Hua, Adikaram Poorni, Pandey Mritunjay, Genis Allison, Simonds William F. Optimization of genome editing through CRISPR-Cas9 engineering. *Bioengineered*. 2016;7(3):166-174. doi:10.1080/21655979.2016.1189039.

52. Doench John G, Fusi Nicolo, Sullender Meagan, Hegde Mudra, Vaimberg Emma W, Donovan Katherine F, Smith Ian, Tothova Zuzana, Wilen Craig, Orchard Robert, Virgin Herbert W, Listgarten Jennifer, Root David E. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*. 2016;34(2):184-191. doi:10.1038/nbt.3437.

53. Ran F Ann, Cong Le, Yan Winston X, Scott David A, Gootenberg Jonathan S, Kriz Andrea J, Zetsche Bernd, Shalem Ophir, Wu Xuebing, Makarova Kira S, Koonin Eugene V, Sharp Phillip A, Zhang Feng. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*. 2015;520:186. http://dx.doi.org/10.1038/nature14299.

54. Tycko Josh, Myer Vic E, Hsu Patrick D. Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. *Mol Cell*. 2016;63(3):355-370. doi:10.1016/j.molcel.2016.07.004.

55. Hsu Patrick D, Scott David A, Weinstein Joshua A, Ran F Ann, Konermann Silvana, Agarwala Vineeta, Li Yinqing, Fine Eli J, Wu Xuebing, Shalem Ophir, Cradick Thomas J, Marraffini Luciano A, Bao Gang, Zhang Feng. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotech*. 2013;31(9):827-832. http://dx.doi.org/10.1038/nbt.2647.

56. Bikard David, Jiang Wenyan, Samai Poulami, Hochschild Ann, Zhang Feng, Marraffini Luciano A. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res*. 2013;41(15):7429-7437. doi:10.1093/nar/gkt520.

57. Nishimasu Hiroshi, Cong Le, Yan Winston X, Ran F Ann, Zetsche Bernd, Li Yinqing, Kurabayashi Arisa, Ishitani Ryuichiro, Zhang Feng, Nureki Osamu. Crystal Structure of *Staphylococcus aureus* Cas9. *Cell*. 2015;162(5):1113-1126. doi:10.1016/j.cell.2015.08.007.

58. Nishimasu Hiroshi, Ran F Ann, Hsu Patrick D, Konermann Silvana, Shehata Soraya I, Dohmae Naoshi, Ishitani Ryuichiro, Zhang Feng, Nureki Osamu. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. 2014;156(5):935-949. doi:10.1016/j.cell.2014.02.001.

59. Singh Vijai, Braddick Darren, Dhar Pawan Kumar. Exploring the potential of genome editing CRISPR-Cas9 technology. *Gene*. 2017;599:1-18. doi:10.1016/j.gene.2016.11.008.

60. Blog for cell transfection and transduction: Genome Editing: the CRISPR/Cas9 system. http://transfections.blogspot.pt/2014/12/genome-editing-crisprcas9-system.html. Accessed December 12, 2017.

61. Kleinstiver Benjamin P, Prew Michelle S, Tsai Shengdar Q, Topkar Ved V., Nguyen Nhu T, Zheng Zongli, Gonzales Andrew PW, Li Zhuyun, Peterson Randall T, Yeh Jing-Ruey Joanna, Aryee Martin J, Joung J Keith. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 2015;523(7561):481-485. doi:10.1038/nature14592.

62. Friedland Ari E, Baral Reshica, Singhal Pankhuri, Loveluck Katherine, Shen Shen, Sanchez Minerva, Marco Eugenio, Gotta Gregory M, Maeder Morgan L, Kennedy Edward M, Kornepati Anand VR, Sousa Alexander, Collins McKensie A, Jayaram Hari, Cullen Bryan R, Bumcrot David. Characterization of *Staphylococcus aureus* Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol*. 2015;16(1):257. doi:10.1186/s13059-015-0817-8.

63. Kleinstiver Benjamin P, Prew Michelle S, Tsai Shengdar Q, Nguyen Nhu T, Topkar Ved V, Zheng Zongli, Joung J Keith. Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol*. 2015;33:1293. http://dx.doi.org/10.1038/nbt.3404.

64. Jiang Wenyan, Bikard David, Cox David, Zhang Feng, Marraffini Luciano A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotech*. 2013;31(3):233-239. http://dx.doi.org/10.1038/nbt.2508.

65. Schmid-Burgk Jonathan L. Disruptive non-disruptive applications of CRISPR/Cas9. *Curr Opin Biotechnol*. 2017;48:203-209. doi:10.1016/j.copbio.2017.06.001.

66. Mali Prashant, Esvelt Kevin M, Church George M. Cas9 as a versatile tool for engineering biology. *Nat Methods*. 2013;10(10):957-963. doi:10.1038/nmeth.2649.

67. Gomaa Ahmed A, Klumpe Heidi E, Luo Michelle L, Selle Kurt, Barrangou Rodolphe,

Beisel Chase L. Programmable Removal of Bacterial Strains by Use of Genome-Targeting CRISPR-Cas Systems. *mBio* . 2014;5(1). doi:10.1128/mBio.00928-13.

68. Qi Lei S, Larson Matthew H, Gilbert Luke A, Doudna Jennifer A, Weissman Jonathan S, Arkin Adam P, Lim Wendell A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013;152(5):1173-1183. doi:10.1016/j.cell.2013.02.022.

69. Qi Lei S, Arkin Adam P. A versatile framework for microbial engineering using synthetic non-coding RNAs. *Nat Rev Micro*. 2014;12(5):341-354. http://dx.doi.org/10.1038/nrmicro3244.

70. Mathew Renjith, Chatterji Dipankar. The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol*. 2006;14(10):450-455. doi:10.1016/j.tim.2006.08.002.

71. Gilbert Luke A, Larson Matthew H, Morsut Leonardo, Liu Zairan, Brar Gloria A, Torres Sandra E, Stern-Ginossar Noam, Brandman Onn, Whitehead Evan H, Doudna Jennifer A, Lim Wendell A, Weissman Jonathan S, Qi Lei S. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*. 2017;154(2):442-451. doi:10.1016/j.cell.2013.06.044.

72. Perez-Pinera Pablo, Kocak D Dewran, Vockley Christopher M, Adler Andrew F, Kabadi Ami M, Polstein Lauren R, Thakore Pratiksha I, Glass Katherine A, Ousterout David G, Leong Kam W, Guilak Farshid, Crawford Gregory E, Reddy Timothy E, Gersbach Charles A. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Meth*. 2013;10(10):973-976. http://dx.doi.org/10.1038/nmeth.2600.

73. Maeder Morgan L, Linder Samantha J, Cascio Vincent M, Fu Yanfang, Ho Quan H, Joung J Keith. CRISPR RNA-guided activation of endogenous human genes. *Nat Meth*. 2013;10(10):977-979. http://dx.doi.org/10.1038/nmeth.2598.

74. Plummer Robert J, Guo Yi, Peng Ying. A CRISPR reimagining: new twists and turns of CRISPR beyond the genome-engineering revolution. *J Cell Biochem*. 2017;(August). doi:10.1002/jcb.26406.

75. Tanenbaum Marvin E, Gilbert Luke A, Qi Lei S, Weissman Jonathan S, Vale Ronald D. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*. 2017;159(3):635-646. doi:10.1016/j.cell.2014.09.039.

76. Kamiyama Daichi, Sekine Sayaka, Barsi-Rhyne Benjamin, Hu Jeffrey, Chen Baohui, Gilbert Luke A, Ishikawa Hiroaki, Leonetti Manuel D, Marshall Wallace F, Weissman Jonathan S, Huang Bo. Versatile protein tagging in cells with split fluorescent protein. 2016;7:11046. http://dx.doi.org/10.1038/ncomms11046.

77. Ma Hanhui, Naseri Ardalan, Reyes-Gutierrez Pablo, Wolfe Scot A, Zhang Shaojie, Pederson Thoru. Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc Natl Acad Sci* . 2015;112(10):3002-3007. doi:10.1073/pnas.1420024112.

78. Chen Janice S, Dagdas Yavuz S, Kleinstiver Benjamin P, Welch Moira M, Harrington Lucas B, Sternberg Samuel H, Joung J Keith, Yildiz Ahmet, Doudna Jennifer A. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *bioRxiv*. 2017:160036. doi:10.1101/160036.

79. Sambrook J, Fristch EF, Maniatis T. Molecular cloning: a laboratory manual, 2nd Ed. *Cold Spring Harb Lab Press Cold Spring Harb New York*. 1989.

80. Kraemer Ginger Rhoads, Iandolo John J. High-frequency transformation of *Staphylococcus aureus* by electroporation. *Curr Microbiol*. 1990;21(6):373-376. doi:10.1007/BF02199440.

81. Oshida T, Tomasz A. Isolation and characterization of a Tn551-autolysis mutant of *Staphylococcus aureus*. *J Bacteriol*. 1992;174(15):4952-4959.

82. Jorge Ana M, Hoiczyk Egbert, Gomes João P, Pinho Mariana G. EzrA contributes to the regulation of cell size in *Staphylococcus aureus*. *PLoS One*. 2011;6(11). doi:10.1371/journal.pone.0027542.

83. Tavares Andreia C, Fernandes Pedro B, Carballido-López Rut, Pinho Mariana G. MreC and MreD proteins are not required for growth of *Staphylococcus aureus*. *PLoS One*. 2015;10(10):1-17. doi:10.1371/journal.pone.0140523.

84. Monteiro João M, Fernandes Pedro B, Vaz Filipa, Pereira Ana R, Tavares Andreia C, Ferreira Maria T, Pereira Pedro M, Veiga Helena, Kuru Erkin, VanNieuwenhze Michael S, Brun Yves V., Filipe Sérgio R, Pinho Mariana G. Cell shape dynamics during the staphylococcal cell cycle. *Nat Commun*. 2015;6:8055. doi:10.1038/ncomms9055.

85. Wu S, de Lencastre H, Sali A, Tomasz A. A phosphoglucomutase-like gene essential for the optimal expression of methicillin resistance in *Staphylococcus aureus*: molecular cloning and DNA sequencing. *MicrobDrug Resist*. 1996;2(2):277-286. doi:10.1089/mdr.1996.2.277.

86. Pinho Mariana G, de Lencastre Herminia, Tomasz Alexander. Transcriptional Analysis of the *Staphylococcus aureus* Penicillin Binding Protein 2 Gene. *J Bacteriol*. 1998;180(23):6077-6081.

87. Nair Dhanalakshmi, Memmi Guido, Hernandez David, Bard Jonathan, Beaume Marie, Gill Steven, Francois Patrice, Cheung Ambrose L. Whole-Genome Sequencing of *Staphylococcus aureus* Strain RN4220, a Key Laboratory Strain Used in Virulence Research, Identifies Mutations That Affect Not Only Virulence Factors but Also the Fitness of the Strain. *J Bacteriol*. 2011;193(9):2332-2335. doi:10.1128/JB.00027-11.

88. Gill Steven R, Fouts Derrick E, Archer Gordon L, Mongodin Emmanuel F, DeBoy Robert T, Ravel Jacques, Paulsen Ian T, Kolonay James F, Brinkac Lauren, Beanan Mauren, Dodson Robert J, Daugherty Sean C, Madupu Ramana, Angiuoli Samuel V, Durkin A Scott, Haft Daniel H, Vamathevan Jessica, Khouri Hoda, Utterback Terry, Lee Chris, Dimitrov George, Jiang Lingxia, Qin Haiying, Weidman Jan, Tran Kevin, Kang Kathy, Hance Ioana R, Nelson Karen E, Fraser Claire M. Insights on Evolution of Virulence and Resistance from the Complete Genome Analysis of an Early Methicillin-Resistant *Staphylococcus aureus* Strain and a Biofilm-Producing Methicillin-Resistant *Staphylococcus epidermidis* Strain. *J Bacteriol*. 2005;187(7):2426-2438. doi:10.1128/JB.187.7.2426-2438.2005.

89. Brakstad OG, Aasbakk K, Maeland JA. Detection of *Staphylococcus aureus* by polymerase chain reaction amplification of the nuc gene. *J Clin Microbiol*. 1992;30(7):1654-1660.

90. Goosens Vivianne J, Monteferrante Carmine G, Van Dijl Jan Maarten. The Tat system of Gram-positive bacteria. *Biochim Biophys Acta - Mol Cell Res*. 2014;1843(8):1698-1706. doi:10.1016/j.bbamcr.2013.10.008.

91. Shaner Nathan C, Lambert Gerard G, Chammas Andrew, Ni Yuhui, Cranfill Paula J, Baird Michelle A, Sell Brittney R, Allen John R, Day Richard N, Israelsson Maria, Davidson Michael W, Wang Jiwu. A bright monomeric green fluorescent protein derived

from *Branchiostoma lanceolatum*. *Nat Methods*. 2013;10(5):407-409. doi:10.1038/nmeth.2413.

92. Pearson Melanie M, Sebaihia Mohammed, Churcher Carol, Quail Michael A, Seshasayee Aswin S, Luscombe Nicholas M, Abdellah Zahra, Arrosmith Claire, Atkin Becky, Chillingworth Tracey, Hauser Heidi, Jagels Kay, Moule Sharon, Mungall Karen, Norbertczak Halina, Rabbinowitsch Ester, Walker Danielle, Whithead Sally, Thomson Nicholas R, Rather Philip N, Parkhill Julian, Mobley Harry LT. Complete genome sequence of uropathogenic *Proteus mirabilis*, a master of both adherence and motility. *J Bacteriol*. 2008;190(11):4027-4037. doi:10.1128/JB.01981-07.

93. Alyahya S Anisah, Alexander Roger, Costa Teresa, Henriques Adriano O, Emonet Thierry, Jacobs-Wagner Christine. RodZ, a component of the bacterial core morphogenic apparatus. *Proc Natl Acad Sci U S A*. 2009;106(4):1239-1244. doi:10.1073/pnas.0810794106.

94. Shcherbo Dmitry, Shemiakina Irina I, Ryabova Anastasiya V, Luker Kathryn E, Schmidt Bradley T, Souslova Ekaterina A, Gorodnicheva Tatiana V, Strukova Lydia, Shidlovskiy Konstantin M, Britanova Olga V, Zaraisky Andrey G, Lukyanov Konstantin A, Loschenov Victor B, Luker Gary D, Chudakov Dmitriy M. Near-infrared fluorescent proteins. *Nat Meth*. 2010;7(10):827-829. http://dx.doi.org/10.1038/nmeth.1501.

95. Zhao Changlong, Shu Xueqin, Sun Baolin. Construction of a Gene Knockdown System Based on Catalytically Inactive ("Dead") Cas9 (dCas9) in *Staphylococcus aureus*. *Appl Environ Microbiol*. 2017;83(12). doi:10.1128/AEM.00291-17.

96. Peters Jason M, Colavin Alexandre, Shi Handuo, Czarny Tomasz L, Larson Matthew H, Wong Spencer, Hawkins John S, Lu Candy HS, Koo Byoung Mo, Marta Elizabeth, Shiver Anthony L, Whitehead Evan H, Weissman Jonathan S, Brown Eric D, Qi Lei S, Huang Kerwyn Casey, Gross Carol A. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell*. 2016;165(6):1493-1506. doi:10.1016/j.cell.2016.05.003.

97. Liu Xue, Gallay Clement, Kjos Morten, Domenech Arnau, Slager Jelle, Van Kessel Sebastiaan P, Knoops Kèvin, Sorg Robin A, Zhang Jing-Ren Jing-Ren, Veening Jan-Willem Jan-Willem. High-throughput CRISPRi phenotyping in *Streptococcus pneumoniae* identifies new essential genes involved in cell wall synthesis and competence development. *Mol Syst Biol*. 2017;13(5):931. doi:10.1101/088336.